INVESTIGACION Y
CIENCIA

Edición en español de SCIENTIFIC AMERICAN



EL OIDO DE LA LECHUZA

Febrero 1982
300 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 65
Febrero 1982



EDITA: Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Guillermo Jenaro Garrido: *Armas láser*; Julio Rodríguez Villanueva: *Navegación magnética en las bacterias*; Manuel Puigcerver: *Júpiter y Saturno*; Joandomènec Ros: *El oído de la lechuza*; Cristóbal Mezquita: *Fibrinógeno y fibrina*; Luis Bou: *Algebra por ordenador*; J. M. García de la Mora: *El principio antrópico*; Luis Bou: *Juegos matemáticos*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

José Campos Gutiérrez

Libros:

Ramón Margalef, Luis Alonso, Agustín Albarracín y F. García de Valdecasas



LA PORTADA

La fotografía de la portada muestra la cara de la lechuza común, *Tyto alba*. El anillo que sobresale de su cabeza es una bobina electromagnética que forma parte del montaje experimental empleado para determinar la precisión con que el ave localiza el sonido y las señales que selecciona de entre las que se hallan presentes en los sonidos naturales. El ave se posa de manera que la bobina cefálica interseque campos magnéticos horizontales y verticales; los cambios en la corriente de la bobina indican movimientos de la cabeza cuando la lechuza localiza un altavoz móvil. Estos ensayos han demostrado que su capacidad para localizar el origen de un sonido es mejor que la de cualquier otra especie de la que se tengan datos. Tal precisión es imprescindible para la caza aérea de ratones de campo durante la noche. La cara de la lechuza es parte fundamental de su sistema auditivo. Los órganos externos son dos conductos que corren a través de las plumas faciales. Asimétricos en el sentido vertical, los oídos informan, a la vez, sobre el ángulo vertical y el ángulo horizontal de una fuente sonora.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:
España:
Un año (12 números): 3.300 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
340 pesetas
Ejemplar atrasado extraordinario:
440 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.º, Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

Fotografía de la portada de Jon Brenneis

Página	Fuente
9-15	Ian Worpole
17	Richard P. Blakemore y Nancy Blakemore
18	Alan D. Iselin
19	Richard P. Blakemore
20	Denise Maratea y Nancy Blakemore (*arriba*), Richard P. Blakemore (*abajo, izquierda*), Richard P. Blakemore y Wendy O'Brien (*abajo, derecha*)
21	David L. Balkwill y Denise Maratea (*arriba*); Richard P. Blakemore y M. Ingerfeld, Universidad de Canterbury (*abajo*)
22-24	Alan D. Iselin
27	José Olivares
30-34	Eulogio J. Bedmar, José Olivares y Miguel Alonso
37-40	José Campos y M. Alonso
43	Jet Propulsion Laboratory, Instituto de Tecnología de California
44-46	Walken Graphics
47	Jet Propulsion Laboratory, Instituto de Tecnología de California y Walken Graphics
48-50	Jet Propulsion Laboratory, Instituto de Tecnología de California
51	Andrew P. Ingersoll
52	Gareth P. Williams
53	Walken Graphics
56	Eric I. Knudsen, Facultad de Medicina de la Universidad de Stanford
58-68	Tom Prentiss
71	Emil O. Bernstein y Eila Kairinen, Gillette Research Institute
72	Robley C. Williams, Universidad de California en Berkeley
73-79	George V. Kelvin
83	Gabor Kiss
84	Carol Hansen
86-90	Gabor Kiss
91	R. F. Bonifield
95-99	Ian Worpole
100	Laszlo Kubinyi
102	Ian Worpole
104-109	Ilil Arbel
110-111	R. F. Bonifield
112-114	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

KOSTA TSIPIS ("Armas láser"), físico del Instituto de Tecnología de Massachusetts, es director adjunto del Programa Científico y Tecnológico para la Seguridad Internacional que se desarrolla en el departamento de física de aquella institución. Nacido en Grecia, se trasladó a los Estados Unidos en 1954 para estudiar física e ingeniería eléctrica. Tras licenciarse por la Universidad Rutgers, en 1958, se doctoró por la de Columbia en 1966. Ese mismo año se incorporaría al departamento de física del MIT.

RICHARD P. BLAKEMORE y RICHARD B. FRANKEL ("Navegación magnética en bacterias") describen su trabajo en común como un "fructífero y saludable mutualismo obligado entre un físico y un biólogo" y como "una verdadera colaboración interdisciplinar en la que cada uno ha tenido que aprender el lenguaje del otro para poder llevarla adelante". Blakemore, el biólogo, descubrió el fenómeno de la magnetotaxia en bacterias hace seis años, cuando todavía era estudiante de segundo ciclo de la Universidad de Massachusetts en Amherst. Tras doctorarse en microbiología ingresó en el claustro docente de la Universidad de New Hampshire, donde enseña esa materia. Frankel, el físico, dice haberse convertido en "empedernido biólogo-adicto" desde que unió sus fuerzas a las de Blakemore. Licenciado por la Universidad de Missouri, se doctoró por la de California en Berkeley. Desde 1965 es miembro de la plantilla de investigadores del Francis Bitter National Magnet Laboratory, adscrito al Instituto de Tecnología de Massachusetts.

EULOGIO BEDMAR y JOSE OLIVARES ("Limitaciones de la fijación biológica de nitrógeno") se hallan adscritos a la Estación Experimental del Zaidín, Granada, centro del Consejo Superior de Investigaciones Científicas. Bedmar se dedica, desde 1974, a la investigación sobre diferentes aspectos de la simbiosis *Rhizobium*-leguminosa. Es doctor en ciencias biológicas por la Universidad de Granada y ha publicado varios trabajos sobre el tema. Olivares se doctoró en farmacia en la misma institución superior. Tras completar su formación en Alemania fue becario de la Fundación March en España. Invitado por distintas universidades y centros de investigación ha dado cursos y conferencias en varios países de habla española. Dirige el departamento de microbiología del Zaidín.

ANDREW P. INGERSOLL ("Júpiter y Saturno") enseña ciencias planetarias en el Instituto de Tecnología de California (Cal Tech). Se licenció por el Amherst College en 1960 y se recibió de doctor por la Universidad de Harvard en 1965. En 1966 ingresó en el Cal Tech. Especialista en atmósferas planetarias, ha tomado parte en el análisis de los datos de varias misiones espaciales norteamericanas. Recientemente, la NASA le concedió la medalla al éxito científico por su trabajo en el proyecto Voyager. Escribe: "Vivo con mi esposa, cinco hijos, algunos estudiantes de segundo ciclo del Cal Tech y un investigador postdoctoral en un hermoso caserón de Pasadena, que a veces ha recibido el nombre de comuna de científicos que trabajan duro. Me divierten los problemas y la competividad; el mérito de mantener mi comportamiento dentro de las formas civilizadas corresponde enteramente a mi familia y amigos".

ERIC I. KNUDSEN ("El oído de la lechuza") es profesor adjunto de neurobiología de la Facultad de Medicina de la Universidad de Stanford. Obtuvo sus grados académicos por la Universidad de California: licenciado en 1971 (en el campus de Santa Bárbara) y doctorado en ciencias en 1976 (en el de San Diego). "Empecé mi carrera de investigación, escribe, interesándome, aún estudiante, por la bioluminiscencia de un alcionario, animal primitivo que pertenece a uno de los tipos más sencillos (Celentéreos). Desde entonces mi interés ha recorrido progresivamente la escala evolutiva, desde la cacerola de las Molucas (tesis de licenciatura) hasta el pez gato (tesis doctoral), y de allí a la lechuza (trabajo postdoctoral)."

RUSSELL F. DOOLITTLE ("Fibrinógeno y fibrina") enseña bioquímica en la Universidad de California en San Diego. Estudió biología en las universidades de Wesleyan y Harvard, recibiéndose de doctor por esta última en 1962. Tras disfrutar de una beca postdoctoral de dos años en el Instituto Karolinska y en la Universidad de Lund, continuó sus trabajos sobre la estructura y evolución del fibrinógeno en San Diego. Durante el curso académico 1971-1972 desarrolló su trabajo en el Wolfson College de la universidad oxoniense.

RICHARD PAVELLE, MICHAEL ROTHSTEIN y JOHN FITCH ("Algebra por ordenador"), aunque distantes geográficamente, colaboran en la aplicación de métodos computarizados a problemas matemáticos de diversas ramas de la ciencia. Pavelle forma parte del Laboratorio de Ciencias de Cómputo del Instituto de Tecnología de Massachusetts. Tras graduarse en ingeniería nuclear por la Universidad de Columbia, se doctoró en matemáticas por la de Sussex. Su "devoción" por el álgebra computarizada nació en 1974, cuando un cálculo, que realizado "a mano" le ocupó tres meses el año anterior, "se rehizo (y confirmó) en dos minutos mediante el sistema MACSYMA del MIT". Rothstein enseña ciencias de cómputo en la Universidad estatal de Kent. Su interés por ellas se suscitó en la Universidad Nacional de Colombia en Bogotá, donde se licenció en matemáticas, en 1970. Una beca Fulbright le llevó después a la Universidad de Wisconsin en Madison, donde se doctoró en 1976. Antes de ingresar como profesor de la estatal de Kent, en 1980, Rothstein enseñó ciencias de cómputo en la Universidad Simón Bolivar de Caracas. Fitch es profesor de ingeniería de soportes lógicos (*software*) en la Facultad de Exactas de la Universidad de Bath, donde ocupa también el cargo de director del centro de cálculo universitario y encabeza el grupo de estudio de análisis numérico y ciencias de cómputo. Se educó en la Universidad de Cambridge, donde se licenció en matemáticas y se recibió de doctor.

GEORGE GALE ("El principio antrópico") es profesor adjunto de filosofía y ciencias físicas de la Universidad de Missouri en Kansas City. Se formó en la Universidad de Santa Clara y en la de California en Davis, donde se doctoró en 1971. "Redacté, cuenta, gran parte de la disertación doctoral siguiendo cursos como estudiante en la Universidad de Oxford. En mi tesis, que versaba sobre 'Leibniz: el físico como filósofo', examiné el influjo de las concepciones físicas leibnizianas sobre su sistema filosófico, así como sobre los orígenes de la dinámica de campos. La historia de los comienzos de la teoría de campos sigue siendo uno de mis principales intereses, centrándose otros en la física moderna, sobre todo en las partículas y en la relatividad... Cuando no estoy enfrascado en algún libro, me ensucio las manos cultivando mi viña."

Hace...

José M.ª López Piñero

... trescientos años

Falleció Juan Caramuel y Lobkowitz, destacada figura de la renovación de las ciencias fisicomatemáticas en la España del siglo XVII.

Había nacido el año 1606, en Madrid, de ascendencia bohemia y flamenca. Su biografía corresponde plenamente al carácter exuberante y poco sereno que tópicamente suele asociarse al período barroco. Después de cursar humanidades en la Universidad de Alcalá, tomó el hábito cisterciense. Completó más tarde su formación en diferentes centros españoles y extranjeros, doctorándose por último en teología en la Universidad de Lovaina. Protegido por el monarca español Felipe IV, por el emperador Fernando III y por el papa Alejandro VII, desempeñó importantes cargos eclesiásticos, hasta ser nombrado obispo de Campagna, en el reino de Nápoles, y luego de Vigevano, en el Milanesado. La mayor parte de su vida la pasó en Bohemia y en distintas localidades de los dominios españoles en Italia. Esto le permitió una independencia y un conocimiento de las nuevas corrientes europeas excepcionales entre los intelectuales y científicos de su época residentes en la misma España. No obstante, su ausencia desde los años de juventud le impidió ejercer una influencia directa sobre la actividad científica en nuestro país. A pesar de su continuada conexión con sus problemas y del "españolismo" de su obra, puesto de relieve por Ceñal, ni su biografía ni su producción científica pueden considerarse con propiedad integradas en la realidad social española de su época. Interesa más bien como un "español fuera de España" cuya posición de adelantado y cuyo prestigio fueron después indirectamente aprovechados por los *novatores* que vivían aquí.

Caramuel no fue un profesional de la ciencia, ni siquiera un clérigo consagrado preferentemente a los estudios científicos, como después lo fueron otros *novatores*, así José de Zaragoza o Tomás Vicente Tosca. Su producción escrita es famosa, ante todo, por el extraordinario número de títulos que comprende y por la diversidad de materias sobre las que trata. Buena parte está dedicada a la teología y a variados temas religiosos o eclesiásticos. Incluye, asimismo, estudios de musicología, de gramática y de estética, textos de carácter político e histórico, además de obras de contenido filosófico y una serie nada escasa de libros científicos. No hace falta subrayar que tan desmesurada fecundidad perjudicó inevitablemente el rigor de las exposiciones. Tampoco le favoreció su afán de originalidad, que le llevó en ocasiones a defender auténticas extravagancias o falsas novedades basadas en débiles fundamentos.

La importancia de la obra científica de Caramuel queda desenfocada si atendemos a los encomiásticos pero interesados elogios que de ella hicieron los innovadores españoles del período siguiente, como Zapata, Sarmiento o Feijóo, que todavía resuenan en los juicios de Menéndez Pelayo. Su interés, en efecto, no reside en supuestas aportaciones creadoras, ni siquiera en su solidez, sino en que estuvo ampliamente abierto a las nuevas tendencias en unas fechas relativamente tempranas. De esta manera, fue uno de los primeros puentes de comunicación entre las corrientes modernas europeas y el ambiente científico español en lo relativo a cuestiones matemáticas, astronómicas y físicas.

Su labor en estas disciplinas dependió estrechamente de Descartes. Ceñal, que ha puesto de relieve esta dependencia en la esfera del pensamiento filosófico, no se ha limitado a recoger su relación directa con el gran filósofo y científico francés, sino que ha subrayado su pertenencia a una "órbita cartesiana" de autores europeos, entre los que se encuentran Gassendi, Maignan y la compleja figura de Atanasio Kircher, de personalidad tan afín a la suya.

La física de Caramuel continuaba siendo filosofía natural, aunque en abierta ruptura con los principios clásicos. Por ello la expuso principalmente en sus escritos filosóficos, que corresponden a un adversario decidido de las doctrinas aristotélicas. "Instituo novam accademiam contra peripateticam" afirma, por ejemplo, en la dedicatoria de su *Rationalis et realis Philosophia* (1642). Se opuso de este modo al hilemorfismo, sumándose en varias obras a la teoría cartesiana de las cualidades de los cuerpos y de los modos de la extensión, así como a algunos puntos de vista de Gassendi. Esta línea no era, por supuesto, la de la física moderna propiamente dicha, pues continuaba siendo tan apriorista y especulativa como la tradicional. La física de Caramuel, como la de su modelo Descartes, tenía una clara simpatía por las nuevas corrientes, principalmente por su tendencia al mecanicismo y a la matematización, pero no formaba parte todavía de una disciplina positiva e independizada de la filosofía. Entre la abundante producción de Caramuel es posible encontrar algún escrito de orientación más propiamente física, como *Sublimium ingeniorum crux* (1642), dedicado al problema de la caída de los graves sobre base experimental. También incluyó numerosos problemas físicos junto a otros filosóficos y teológicos en su *Mathesis audax* (1642), pero no llegó nunca a ofrecer una exposición sistemática de esta disciplina dentro de un marco no filosófico sino matemático, como después hizo Tosca.

Más interesante desde el punto de vista puramente científico es su producción astronómica. Aunque no fue un astrónomo práctico como Vicente Mut o José de Zaragoza, se le deben interesantes exposiciones teóricas, contenidas en sus libros filosóficos, teológicos o científicos y también en algunos estudios monográficos. Uno de los primeros fue el titulado *Caelestes Maetamorphoses* (1639) o transformaciones de las órbitas circulares teóricas de los planetas. "No es pequeña gloria para Caramuel –afirma Cotarelo de este texto– hacerse cargo, aunque sea para rechazarla, de la primera ley de Kepler, generalmente inadvertida entonces." En otro estudio, *De novem sideribus circa Jovem Visis* (1643), defendió al capuchino Reitha que creyó haber descubierto cinco satélites nuevos en torno a Júpiter. Pero la más importante de sus aportaciones en este terreno es la contenida en la cuarta parte de su *Cursus Mathematicus*, que lleva el expresi-

vo título de *Mathesis Astronomica in Physicae Tribunali damnata* (1668). Lo que no había llegado a hacer con la física lo hizo con la astronomía, que aparece aquí como una ciencia positiva formulada en lenguaje matemático, es decir, en la auténtica línea de la "mecánica celeste".

Numeri naturales.	Logarithmi Vlacquii.	Logarith. Neperi, & Kepleri.	Logar. Caramuelis.
1	0.	2302,585.20.	10.
10	1.	2072,326.68.	9.
100	2.	1842,068.16.	8.
1,000	3.	1611,809.64.	7.
10,000	4.	1381,551.12.	6.
100,000	5.	1151,292.60.	5.
1;000,000	6.	921,034.08.	4.
10;000,000	7.	690,775.56.	3.
100;000,000	8.	460,517.04.	2.
1,000;000,000	9.	230,258.52.	1.
10,000;000,000	10.	000,0000.00.	0.

Comparación entre los logaritmos de Ulacq, Neper y los "perfectos" de Caramuel, según el **Cursus Mathematicus** *(1667-68), de este último*

La parte más valiosa de la obra científica de Caramuel es la matemática, tal como ha puesto de relieve el excelente estudio que le ha dedicado Santiago Garma. Su obra fundamental en este terreno es el *Cursus Mathematicus* (1667-68), ambicioso tratado en cuatro partes, de las cuales las dos primeras –*Mathesis Vetus novis operationum compendiis et demonstrationibus dilucidata* y *Mathesis Nova, veterum inventis confirmata*– son de contenido propiamente matemático; la segunda –*Mathesis Arquitectonica*– está dedicada a la arquitectura y a la técnica de la construcción; y la tercera, como ya hemos dicho, a la astronomía. Aparecida en 1667-68 la primera edición, en 1670 se publicó una segunda de las dos primeras partes, con el título de *Mathesis biceps, vetus et nova*. También de la tercera parte apareció otra edición, traducida al castellano por el propio Caramuel, que la tituló *Architectura civil recta, y obliqua* (1678). Resulta muy significativo que la dedicara al príncipe Juan José de Austria. También lo es que justificara el uso del castellano en los siguientes términos: "Escribo en mi lengua materna por ver que grandes ingenios en España se ocupan, por no vivir ociosos, en escribir novelas, comedias o diversos poemas... si yo consigo que los que son aplaudidos en los teatros de España por el ingenio y elocuencia con que sus comedias escriben dejen de ser pródigos de su tiempo y talento, y se apliquen a leer este pequeño libro, quedaré muy gozoso; seguro que si lo hicieren, conocerán que me he ocupado de su servicio, y que tienen obligación de confesar que los rasgos de mi pluma y pincel les ha sido útiles y provechosos. Escribo claro, porque tomo la pluma para escribir lo que todos entiendan y no para que estudien y disputen en qué sentido debe ser explicado... si la misma gramática no tolera que la oscurezcan con retóricas flores, menos lo tolera la arquitectura, que de suyo es oscura, por fundarse en cuentas aritméticas y geométricas demostraciones".

Garma destaca, entre las contribuciones matemáticas de Caramuel, su estudio completo de los sistemas de numeración, la exposición de la teoría del número entero en forma analítica, el estudio detallado y moderno de la combinatoria y el planteamiento y resolución de los problemas iniciales del cálculo de probabilidades. Como ejemplo, véase su forma de matematizar los juegos de azar:

"Si jugamos con dados comunes, que tienen seis lados, y alguien apuesta una moneda de oro, diciendo que *saldrá un seis y ningún otro número*, si no sale un seis, deberá pagar una moneda de oro, y si sale, recibir cinco. Esto es válido si usamos un solo dado.

"¿Qué sucede si usamos dos o tres dados? Si alguien con dos dados dice que *van a salir 12 puntos*, tiene un solo modo de acertar y 35 de equivocarse. Por lo tanto, si pone en peligro una moneda de oro, debe esperar 35. Considera la tabla siguiente:

Número de dados		Esperanza	Peligro	Por lo tanto, pagas	Por lo tanto, recibes
2	12	1	35	1	35
3	11	2	34	1	17
4	10	3	33	1	11
5	9	4	32	1	8
6	8	5	31	1	$6^{1/5}$
7	7	6	30	1	5

"Tienes a la vista seis columnas. En la primera y la segunda están los números de los dados que te propones echar. En la tercera, la esperanza, es decir, de cuántos modos puedes conseguir lo apostado. En la cuarta, el peligro, o sea, de cuántos modos puedes no conseguirlo. En la quinta, cuánto debes pagar. En la última cuánto debes recibir cada vez que salga el número apostado, mientras que pierdes una moneda de oro cada vez que no salga.

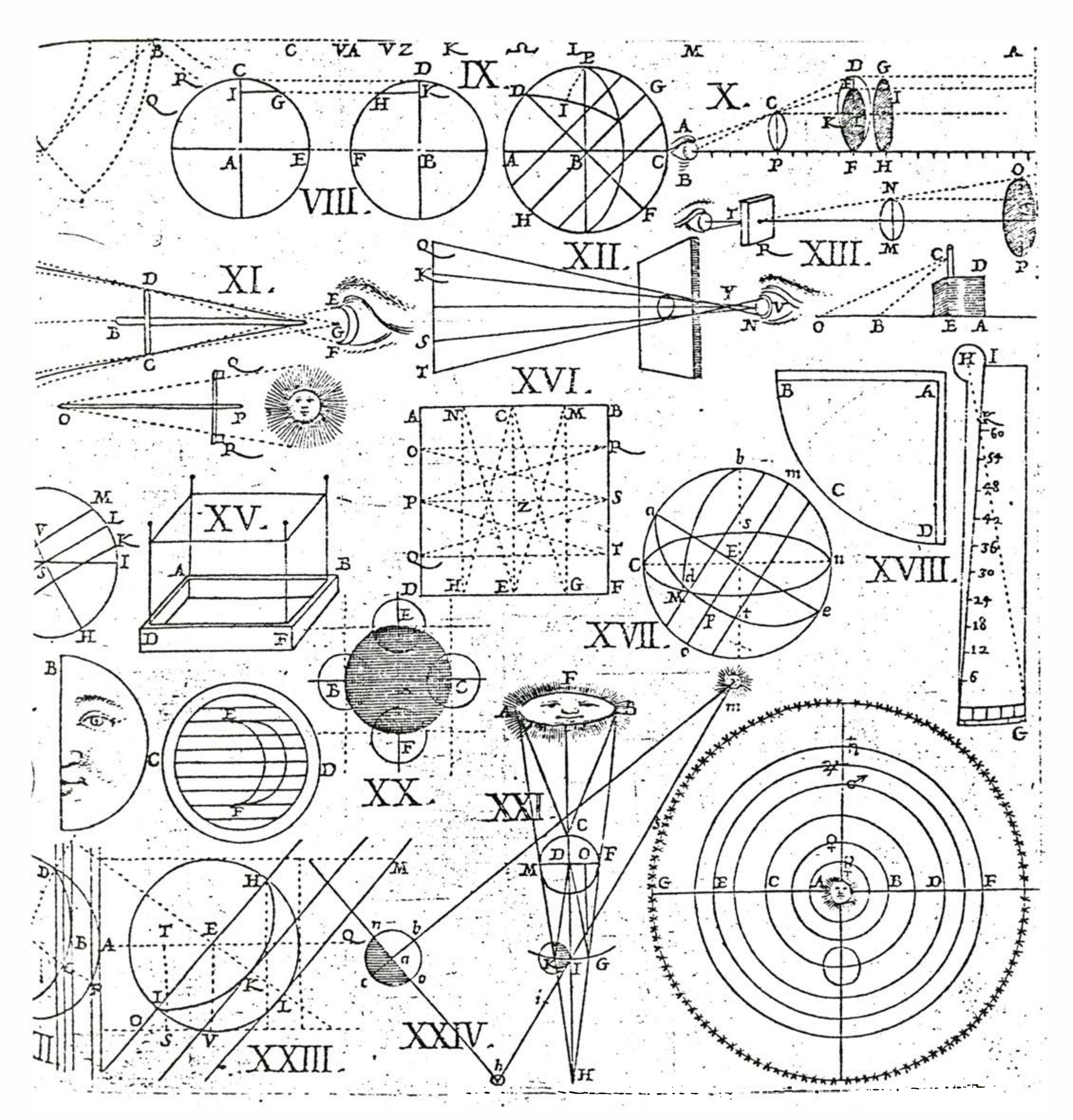
"De aquí puede deducirse lo que sucede si se utilizan tres o más dados. El que tiene que esperar una puntuación determinada siempre se expone a un peligro mayor que el que apuesta de forma disyuntiva por una u otra."

Caramuel publicó también las primeras tablas de logaritmos editadas por un científico español e ideó un sistema que denominó de "logaritmos perfectos". Aunque ya José de Zaragoza demostró que no representaba ninguna ventaja sobre los logaritmos de Briggs, esta invención coloca a Caramuel, en opinión de Sánchez Pérez, entre los antecesores del cálculo de los logaritmos. En relación con este tema, Caramuel formuló con claridad la idea de progreso científico, como poco después harían también otros *novatores* españoles:

"Es cuestión muy reñida, promovida más por prejuicios de ingenios pusilánimes que gobernada por razón, la que pregunta: ¿Qué es lo que supieron los antiguos? Si lo supieron todo y tan todo, ¿qué no dejaron que buscar o cosa de más a los modernos? Si aquéllos se hubiesen empeñado en defender la afirmación, tuviera excusa su arrogancia; pero que éstos sean tan enemigos de sí mismos que, contra toda razón y experiencia, se quieran condenar y abatir, me parece cosa maravillosa. No es mi intento quitar o cercenar a los antiguos las alabanzas que merecen; supieron mucho, mas no todo, y hoy los moder-

nos, con haberse descubierto nuevas ciencias, siempre dejan que saber a la posteridad. Estaban sosegados los cosmógrafos, las partes del mundo eran tres: Europa, Africa y Asia; cuanto se decía de los antípodas se oía con risa, como cosa imposible. Era común sentencia que no se podía navegar al Océano, ya que no obedecían nuestras expediciones al imperio de Hércules que las puso término, diciendo NON PLUS ULTRA, *No hay más a donde se puede andar o navegar*... Salió Colón, navegó y, descubriendo grandes y ricas islas, probó que la cosmografía antigua con tiranía nos había encarcelado en los términos del NON PLUS ULTRA que en España puso la Antigüedad. PLUS ULTRA, a pesar de Hércules, dijo Colón y salió con la suya...

"Del mismo modo sucedió en el orbe literario... en nuestro tiempo los que profesan letras con generosa audacia derribaron las columnas que habían erigido los Hércules antiguos y borraron la inscripción que decía NON PLUS ULTRA, *No hay ya más que saber* y, engolfándose en el Océano de la Matemática, descubrieron una India ultramarina que hoy se llama *Logarítmica*."



Esquemas procedentes de una de las láminas de la **Mathesis biceps, vetus et nova** *(1670) de Caramuel. Entre ellos se encuentra uno correspondiente al sistema heliocéntrico de Copérnico*

Armas láser

¿Podrían los láseres instalados en satélites defender una nación contra el ataque de misiles? No parece fácil que así sea. A los obstáculos técnicos hay que añadir la fragilidad de esas armas ante sencillas contramedidas

Kosta Tsipis

Un láser de alta energía puede taladrar fácilmente, por combustión, una placa metálica de bastante espesor. La industria lo utiliza como herramienta de corte. Capacidad que indujo a pensar, en obvio razonamiento, en su posible uso como arma en caso de guerra. Pero algunos ingenieros de armamento han querido ver en él algo así como la pistola de rayos de las películas de ciencia-ficción. Se apuntaría el rayo láser contra un avión, un misil atacante o cualquier otro objetivo, desintegrándolo al instante. Y puesto que se propaga con la velocidad de la luz, no cabría correr más que él o esquivarlo. La verdad es que en ensayos preliminares, bajo condiciones controladas, los láseres han destruido pequeños aviones con guía remota.

Recientemente, ha habido quienes en el Congreso de los Estados Unidos, en el Ministerio de Defensa y en la industria aeroespacial han sostenido que los láseres de alta energía tenían capacidad potencial para destruir, en vuelo, misiles balísticos intercontinentales. Al argüir que la Unión Soviética ha realizado un gran esfuerzo para desarrollarlos como armas contra misiles, con el consiguiente retraso norteamericano en este campo, urgen a la Administración Reagan a que amplíe generosamente el programa nacional de armas láser, que ya recibe en la actualidad 300 millones de dólares al año.

¿Con qué objetivos? Sería el principal desplegar en órbitas terrestres, a lo largo de los próximos 10 años, una red de armas láser de gran potencia, que tendrían por misión destruir los misiles intercontinentales rusos, en cuanto fueran lanzados. Habría que procurar también desarrollar esos ingenios para su lanzamiento desde tierra en misión de ataque contra satélites enemigos y de defensa de objetivos vitales, de la agresión posible de aviones o misiles tácticos.

Los efectos de tales armas hipotéticas sobre el equilibrio militar mundial y sobre las expectativas de limitación de armas nucleares debieran merecer, algún día, atenta consideración. De momento, sin embargo, habrá que plantearse una cuestión más apremiante ¿Resulta factible, desde el punto de vista técnico, un sistema eficaz de armas láser? Sostendré aquí que los objetivos fijados para el desarrollo de armas láser tardarán posiblemente en alcanzarse más de los diez años marcados. La verdad es que mientras no se superen bastantes obstáculos fundamentales para el empleo del láser como arma, esas metas no se lograrán nunca. Hay dificultades atribuibles a la física de la propagación de un rayo láser sobre grandes distancias; otras tienen un origen tecnológico o económico.

La capacidad bélica de los láseres ha sido objeto de examen en una serie de seminarios organizados por el Programa de Ciencia y Tecnología para la Seguridad Internacional, del Departamento de Física del Instituto de Tecnología de Massachusets (MIT). Tomaron parte colegas míos del MIT, y yo mismo, así como investigadores procedentes de otras universidades, de la industria y de los laboratorios nacionales de armamento. Hemos llegado a la conclusión siguiente: el láser, como arma defensiva práctica y atendiendo a la adecuada relación coste/eficacia, tiene muy pocas, o ninguna, probabilidades de éxito.

Un láser genera un intenso chorro de ondas electromagnéticas, dotadas todas ellas de idéntica frecuencia, fase y dirección de movimiento; se dice entonces que las ondas son coherentes. La propiedad de la coherencia resulta esencial en su aplicación bélica, puesto que, para ser eficaz en su ataque, un haz de luz láser ha de ser intenso, bien colimado y hallarse en fase las ondas que lo constituyen. Aunque, en principio, la intensidad de la luz de un láser es ilimitada, en la práctica depende de su tamaño y de las propiedades del material donde se genera la luz coherente.

El medio activo de un láser puede ser un sólido, un líquido o un gas. Los láseres de alta energía suelen utilizar un gas molecular. Para iniciar la acción láser debe suministrarse energía externa a las moléculas del gas. Parte de esta energía incrementa la energía cinética de las moléculas y, por tanto, se limita a calentar el gas; parte se absorbe en los movimientos internos de vibración y rotación de las moléculas. La molécula así excitada abandona su estado inferior de energía de vibración o rotación (el estado fundamental) y pasa a ocupar otro más alto. Resulta de ello que los estados de baja energía se despueblan y un número significativo de moléculas ingresa en un estado excitado. Situación que los físicos denominan inversión de la población de moléculas del gas.

La molécula que se halle en un estado excitado retornará a un estado de inferior energía emitiendo un fotón, un cuanto de radiación electromagnética. La frecuencia de la radiación viene determinada enteramente por la diferencia de energía de los dos estados. El funcionamiento del láser depende de una peculiaridad de la interacción de los fotones con la materia. Cuando un fotón emitido por una molécula excitada choca con otra en el mismo estado de excitación, puede estimular la emisión por esta segunda molécula de un fotón adicional de la misma fase y frecuencia que el fotón estimulante. Ambos fotones pueden luego estimular emisiones similares de otras moléculas, de suerte que el número de fotones idénticos en movimiento a través del colectivo de moléculas crece exponencialmente. La naturaleza de este proceso se recoge en la palabra LASER, acuñada como acrónimo de "*L*ight *A*mplification by *S*timulated

*E*mission of *R*adiation" ("amplificación de la luz por emisión estimulada de radiación").

En un láser, un conjunto de moléculas sometido a una inversión de su población se encierra en una "cavidad" óptica con superficies especulares y paralelas en los extremos. Los fotones emitidos por las moléculas excitadas, al ser reflejados por los dos espejos, recorren una y otra vez el medio generador del láser y estimulan la emisión de fotones por nuevas moléculas, pordadores todos de la misma frecuencia y fase. Puesto que sólo los fotones que siguen un camino exactamente perpendicular a los espejos permanecen en la cavidad tiempo suficiente para ser amplificados, la luz generada en su interior constituye un haz muy bien colimado. Se permite que una pequeña porción del haz abandone la cavidad al hacer parcialmente transparente uno de los espejos.

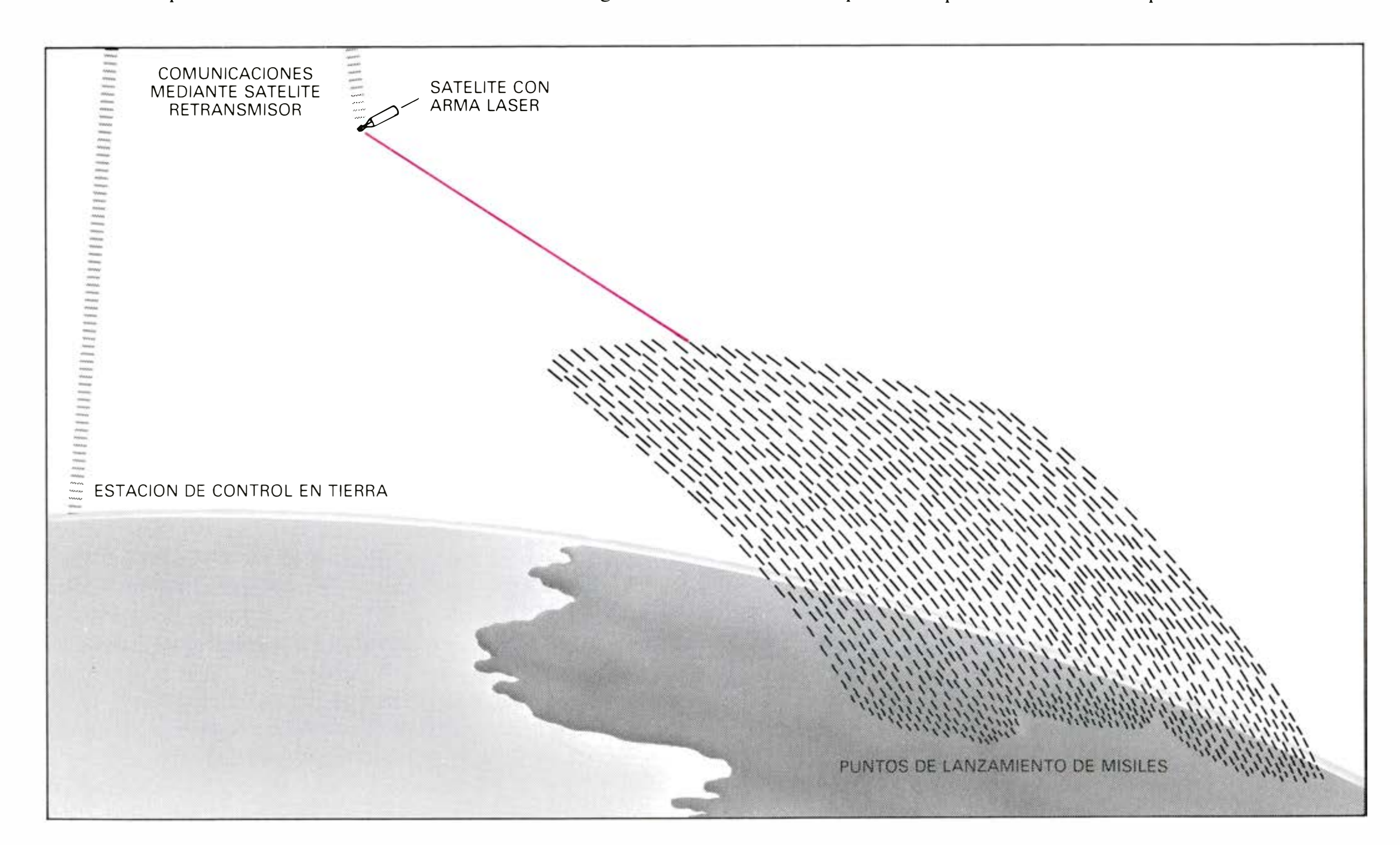
Los fotones cuya frecuencia corresponde a la región visible del espectro electromagnético tienen poca energía: menos de 10^{-19} joule, o watt por segundo. Ahora bien, la emisión de energía de un láser puede alcanzar muchos miles de joule, emitidos en un tiempo sumamente breve, a veces de sólo millonésimas de segundo. ¿A qué obedece ello? A que cabe estimular un número de átomos ingentes (hasta 10^{23}), para que radien muchas veces durante este corto intervalo.

Tres clases de láser de alta energía entran en la consideración de armas potenciales. Se ordenan conforme al mecanismo que crea la inversión de la población en el medio de trabajo. En los láseres por dinámica gaseosa, se generan por combustión gases del estilo del dióxido de carbono. El gas se forma a alta temperatura, lo que comporta que la mayoría de las moléculas se hallen en estados excitados. El gas se enfría repentinamente luego por expansión a través de una serie de toberas; el enfriamiento es tan rápido que las moléculas que ocupan estados excitados no tienen tiempo de volver al estado fundamental. Se crea de este modo una inversión de la población, que origina la emisión de radiación láser inmediatamente después de la expansión.

En un láser de descarga de electrones, se logra la inversión de la población merced a un haz de electrones dirigido a través del medio gaseoso activo. Los electrones ceden parte de su energía en colisiones con las moléculas del gas, originando transiciones a estados energéticos de vibración o rotación más elevados. A través de este mecanismo se prolonga indefinidamente la inversión de la población.

En un láser químico, dos elementos o compuestos químicos se combinan para formar moléculas de un nuevo compuesto; hidrógeno y flúor gaseosos, por ejemplo, pueden combinarse para formar fluoruro de hidrógeno. Las moléculas se originan en un estado excitado; controlando su entorno, podemos lograr la emisión estimulada de radiación antes de que vuelvan al estado fundamental por disipación de su energía en forma de calor.

Estos tres métodos consiguen la inversión de la población, con buen rendimiento, en un nutrido conjunto de moléculas. Pero la práctica impone límites al tamaño de una cavidad óptica eficiente para un láser, así como a la potencia que puede manejar. Gran parte de la investigación y desarrollo empeñados en el campo de los láseres de



ENTRE LAS MISIONES PROPUESTAS para las armas láser está la de defender a una nación contra los misiles balísticos intercontinentales. Se montaría el láser en un satélite en órbita que girase a más de 1000 kilómetros de la superficie de la Tierra, y atacaría a los misiles en su fase de propulsión o despegue, que viene a durar unos ocho minutos. Se necesitarían unos 50 satélites para garantizar que uno al menos estuviera siempre a una distancia eficaz de tiro de las rampas de lanzamiento. A cada satélite le correspondería un lote de hasta 1000 misiles a abatir en ocho minutos, ya que los satélites restantes estarían sobre algún otro lugar de la Tierra. Para los trabajos de detección y seguimiento, se servirían de sensores de infrarrojos o radares. Otros sensores distintos determinarían el error de puntería, o evaluarían el daño causado en la hipótesis de que el haz alcanzara al misil. Un satélite retransmisor, que girase en una órbita más alta, actuaría de enlace de comunicaciones con tierra. No se muestra aquí ninguna de las contramedidas que anulan el arma láser, pero incluso bajo estas condiciones simplificadas las limitaciones físicas (atmosféricas y otras) y tecnológicas hacen del láser un arma inviable.

alta energía tiene por objeto relajar tales restricciones.

Un arma láser diferiría en tres aspectos principales de todos los artefactos desplegados hasta ahora. En primer lugar, transportaría la energía destructora hasta el objetivo en forma de un intenso haz de ondas electromagnéticas, en vez de hacerlo en forma de carga explosiva instalada en la cabeza de un misil o de un proyectil. En segundo lugar, avanzaría con la velocidad de la luz, a unos trescientos millones de metros por segundo; recordemos, a modo de comparación, que un misil supersónico alcanza una velocidad entre 1000 y 2000 metros por segundo. Por último, el haz láser ha de dar caza real al objetivo para que pueda inutilizarlo, mientras que una cabeza nuclear explosiva puede resultar eficaz a bastante distancia del blanco. En consecuencia, para que un arma láser cumpla su cometido, ha de conocerse la posición del objetivo dentro de un margen de error igual a la menor de sus dimensiones, exigiéndose además idéntica precisión al apuntar el láser.

En razón de esas características señaladas, cabría asignar tres clases de misiones a nuestro ingenio. Instalado en un satélite en órbita alrededor de la Tierra podría atacar a los misiles balísticos intercontinentales en fase de lanzamiento, intervalo que viene a durar unos ocho minutos, y podría atacar a los satélites enemigos en órbita. Desde su base en tierra podría abatir satélites y aviones enemigos que lo sobrevolaran; montado en un buque, podría defenderlo de misiles agresores. Por último, un arma láser transportada en avión podría acabar con misiles o aviones enemigos.

En cualquiera de estas misiones, el sistema de arma láser tendría que consumar con éxito una secuencia entera de operaciones. En primer lugar, detectar el blanco y distinguirlo de posibles señuelos y otros objetos de su entorno. Apuntar luego el haz láser al objetivo, seguir sus movimientos y disparar el haz a través del medio que lo rodea. Tras cada disparo habría que determinar si se ha hecho o no blanco. Supongamos un golpe fallido. El sistema tendría que precisar por cuánto, en qué dirección se orientó mal el haz luminoso, corregir la puntería y disparar de nuevo. Después de un impacto, sería preciso determinar si se ha destruido o no el objetivo; en caso negativo, se repitiría el proceso de puntería y disparo. Finalmente, el sistema ha de comunicar al puesto central de tiro los resultados y encaminarse hacia un nuevo blanco si fuere necesario.

Para llevar a cabo todo esto, el sistema necesitará equipo complementario del propio láser. Desde un gran espejo, gobernado por un sistema de mando sumamente preciso para apuntar el haz al objetivo, hasta un conjunto de sensores capaces de detectar, identificar y determinar la posición del objetivo con la precisión y estabilidad requeridas, pasando, además, por precisos dispositivos de control que transmitan las indicaciones de los sensores al espejo de puntería. Sin olvidar aparatos especializados que evalúen el daño causado al objetivo o, en su caso, determinar el error de puntería; amén de obvios equipos generadores de potencia y almacenadores de energía, así como un dispositivo para suministrar energía al láser, en pulsos intensos, en el momento oportuno.

Un sistema de arma láser que opere en el espacio exterior y ataque objetivos situados allende la atmósfera terrestre podría enviar sus rayos a muy largas distancias, porque la luz se propaga sin impedimentos en el vacío. Sin embargo, el haz divergiría algo a causa de la difracción, consecuencia de la naturaleza ondulatoria de la luz. Si se supone que el espejo posee una forma y superficie perfectas, el ángulo por el cual diverge el haz es inversamente proporcional a la frecuencia de la luz del láser y al diámetro del espejo. Puesto que se desea mantener reducida al mínimo la dispersión del haz, un láser con misión bélica funcionaría mejor si generara luz de alta frecuencia y estuviere equipado con un espejo de puntería de gran tamaño.

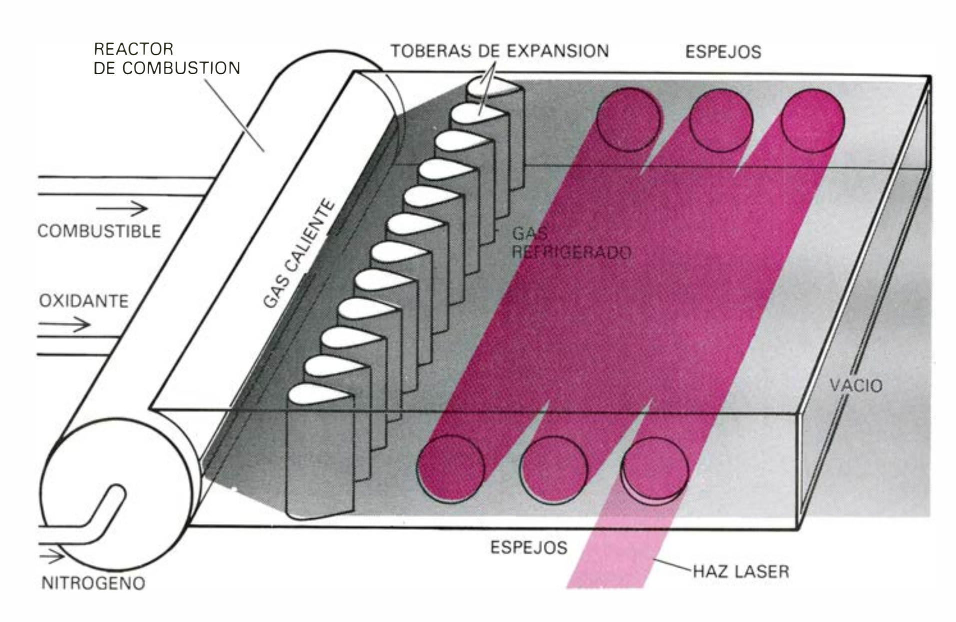
Por otro lado, al atravesar la atmósfera se atenúa y dispersa por varios motivos. Las moléculas del aire y las partículas de materia que contiene (polvo, gotículas de agua, humos) dispersan y absorben la luz. Un haz de infrarrojos generado por un láser de dióxido de carbono vería reducida su intensidad a la mitad tras cuatro kilómetros de recorrido en aire frío y seco, o en 1,5 kilómetros en aire caliente y húmedo. Las nubes, el humo, el polvo, la niebla y la espesa bruma absorberían el haz en su casi totalidad. La eficacia de un arma láser que opere en la atmósfera, resumiríamos, dependerá de las condiciones meteorológicas. Tal dependencia constituye un serio inconveniente en cualquier arma, pero resulta de extrema gravedad en un arma defensiva, que habría de reaccionar ante un ataque desplegado en el momento (y, por tanto, en las condiciones meteorológicas) elegido por el enemigo. Hasta en tiempo despejado el láser puede ser refractado, dispersado o totalmente inte-

DE LAS TRES CLASES de láser consideradas armas potenciales, una es el láser químico. El medio activo es un gas en el que gran parte de sus moléculas han sido transportadas a un estado excitado, creando una "inversión de la población" de las mismas. La luz se amplifica cuando atraviesa repetidamente el medio en la "cavidad" óptica formada por dos espejos. La reacción química entre los dos gases crea el medio activo. El flúor y el hidrógeno que fluyen de una batería de toberas se combinan para originar moléculas de fluoruro de hidrógeno en estado excitado (HF*). Para prolongar la existencia de las moléculas excitadas se mantiene baja la densidad mediante la extracción de los productos de combustión por un sistema de vacío. Cuando las moléculas excitadas (HF*) vuelven después a su estado fundamental (HF), emiten luz.

rrumpido por los fenómenos atmosféricos. La turbulencia origina rápidos cambios locales en la densidad del aire que refractan un rayo de luz o lo hacen divergir. Tenemos un ejemplo de este fenómeno en el parpadeo de las estrellas y de las luces distantes.

Una fracción considerable de la energía de un rayo láser queda absorbida por la atmósfera. De ahí que el aire que va cruzando el haz se caliente y expansione creando un canal de aire de baja densidad. Las ondas luminosas se refractan y alejan de las regiones más calientes y menos densas del medio, lo que determina que el haz diverja. Se ha denominado a este fenómeno "divergencia térmica", y por él se explica el desenfocado y dispersión de un rayo láser en el aire.

Citemos un último obstáculo a la propagación de un haz láser a través de la atmósfera: el riesgo de creación de un plasma. Por ser las ondas luminosas una forma de radiación electromagnética, el haz de luz intenso irá acompañado de un intenso campo eléctrico. A intensidades del orden de los 10 millones de watt por centímetro cuadrado (el valor exacto dependerá de la frecuencia de la radiación), el campo, más fuerte, separa los electrones de los átomos del aire; y así ioniza el medio, originando un plasma. Este absorbe el haz luminoso e interrumpe su transmisión. En virtud de ese efecto se establece, pues, un límite superior para la intensidad del haz láser que se propaga a través de la atmósfera.



LASER POR DINAMICA GASEOSA. Aquí se logra la inversión de la población generando un gas caliente, que se refrigera luego repentinamente. Se quema el combustible con un oxidante para formar dióxido de carbono a alta temperatura. La emisión de luz en el dióxido de carbono se origina en las transiciones entre dos estados excitados. En el gas caliente, muchas moléculas ocupan estados excitados, pero no cabe la acción láser, ya que el nivel de energía láser de la población superior no es mucho mayor que el de la población de nivel inferior; a alta temperatura las moléculas del nivel inferior no lo abandonan por emisión en cascada hasta el nivel fundamental. Cuando el gas se refrigera rápidamente por expansión a través de las toberas, se torna accesible el estado fundamental, así como las transiciones desde el nivel superior al inferior. El nitrógeno añadido a la mezcla transfiere energía a las moléculas de CO_2.

El láser destruiría el blanco por recalentamiento: concentrando en él más energía térmica de la que puede soportar su funcionamiento. Pero los daños infligidos resultarían sólo de la energía absorbida por la superficie del objetivo. Quiero decir que si fuera el blanco de aluminio brillante, absorbería tan sólo el 4 por ciento de la radiación procedente de un láser de infrarojo que lo alcanzara. El resto se reflejaría sin causar desperfecto alguno.

¿Cuánta energía láser absorbería el objetivo? Dependería de la frecuencia de la radiación, de su material de fabricación y del estado de su superficie. Las radiaciones visible e infrarroja se reflejan, poco menos que enteramente, por una superficie metálica pulida; no llegaría en este caso ni al 10 por ciento de la energía transportada por el láser la que se absorbiera, y cumpliera, por tanto, con su misión destructora. Mucho más elevada es la fracción de radiación ultravioleta absorbida por una superficie metálica; aquí, más de la mitad de la energía en forma de radiación ultravioleta que alcanzase el objetivo atentaría contra la integridad de éste.

Hay varios mecanismos por los que el sobrecalentamiento podría destruir o inutilizar un misil u otro blanco parecido. La cantidad de energía por unidad de área que habría de serle transferida con dicho fin dependería del mecanismo elegido y de la vulnerabilidad del objetivo. Por ejemplo, los circuitos electrónicos de un satélite no protegido acabarían por estropearse si el vehículo se hallara sometido a iluminación incesante, a lo largo de varios minutos, por un haz láser con una intensidad de aproximadamente un watt por centímetro cuadrado. Ello viene a suponer 10 veces la intensidad de la luz solar en las capas superiores de la atmósfera. La absorción de 1000 watt por centímetro cuadrado durante un segundo (una cantidad total de energía absorbida de 1000 joule por centímetro cuadrado) fundiría una superficie metálica de milímetros de espesor. Sin embargo, para transferir tal cantidad de energía, un láser de infrarrojo tendría que inyectar en el objetivo unos 20.000 joule por centímetro cuadrado, puesto que la mayor parte de la energía se reflejaría.

Un láser que transmita su energía en breves y potentes destellos o pulsos puede alcanzar una intensidad instantánea de un millón de watt por centímetro cuadrado, si bien la potencia promedio sería muy inferior. La superficie de un objetivo alcanzado por tales pulsos perdería rápidamente su brillo; la fracción de energía del haz absorbida se incrementaría con cada pulso. Cabe pues, en principio, abrir una brecha en el blanco a abatir con el fuego de un haz láser pulsante.

Cuando el objetivo esté en un punto de la atmósfera, una intensidad de unos 10 millones de watt por centímetro cuadrado originaría la ionización del aire inmediato al blanco, creando una capa de plasma donde el haz golpee la superficie. El plasma absorbería la energía del haz, se calentaría hasta la incandescencia (unos 6000 grados Celsius) y se liberaría de aquella energía de dos maneras: emitiendo radiación ultravioleta y expansionándose por vía explosiva. Uno y otro mecanismo aumentarían la proporción de energía del haz cedida al objetivo en un 30 por ciento, aproximadamente, reduciendo por consiguiente, en la misma proporción, la cantidad de energía que el láser tendría que generar.

Un haz pulsante de intensidad extrema llegaría a evaporar el metal que recubriera el blanco. El metal evaporado escaparía de la superficie a alta velocidad y su cantidad de movimiento se equilibraría con una cantidad de movimiento idéntica y de signo contrario, aplicada sobre el objetivo. El impulso generado de esta manera podría despedazar o agrietar un ingenio metálico.

El conocimiento físico de estos efec-

tos nos permite conjeturar, con buena aproximación, las condiciones que habría de reunir un arma láser para llevar a cabo determinadas misiones. Creo que vale la pena detenerse en una misión específica: la de un arma láser en órbita terrestre, diseñada para destruir misiles intercontinentales balísticos del enemigo en fase de propulsión. Aunque se trata de la aplicación más remota, si nos atenemos a su viabilidad y tiempo de desarrollo, resulta la más atractiva desde el punto de vista conceptual, y la más citada cuando se debate en público este sector del armamento.

Los láseres de defensa contra misiles estarían desplegados sobre satélites, en órbitas a unos 100 kilómetros de la superficie terrestre. A tamaña altura, los puntos de lanzamiento ubicados en la Unión Soviética, sólo se hallarían dentro de la distancia eficaz de un satélite breves intervalos de tiempo en cada órbita. Para garantizar que, permanentemente, hubiere al menos un satélite a distancia de tiro, la red tendría que contar con unos cincuenta satélites. Cada satélite habría de ser capaz de destruir, por sí mismo, una fuerza total de hasta un millar de misiles balísticos en fase de propulsión, que viene a durar unos ocho minutos. Por tanto, el satélite sólo podría dedicarle a cada misil algo así como medio segundo.

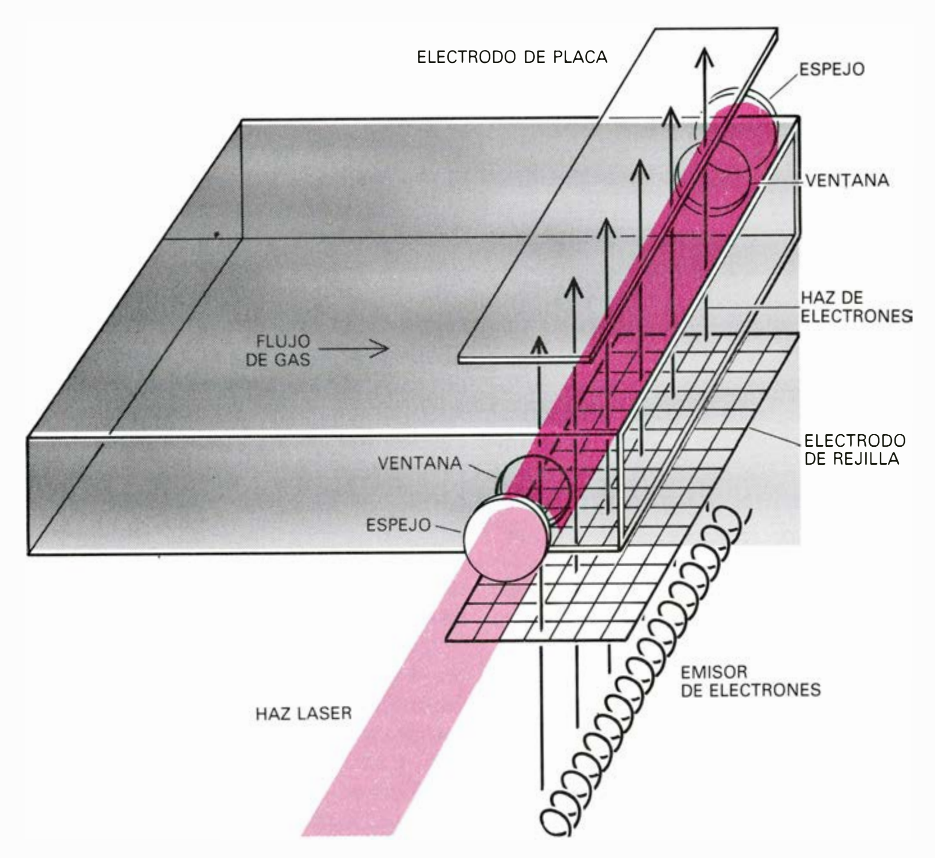
Con mis colegas he hallado un método eficaz para dejar fuera de combate misiles balísticos: agrietar su superficie mediante una carga impulsiva. El agrietamiento resultaría de la absorción de varios pulsos breves, de 1000 joule por centímetro cuadrado cada uno. Aportaría la energía un haz de una intensidad en el blanco que rondaría el millón de watt por centímetro cuadrado y una duración de los destellos de cienmillonésimas de segundo. Las pruebas de laboratorio cifran en torno a diez el número de destellos necesarios para abrir un boquete en el misil.

¿Cuánta energía habría de desarrollar ese láser? Supongamos que el arma en cuestión es un láser de destellos de fluoruro de hidrógeno y que el espejo director de su haz luminoso posee todas las garantías ópticas y mide un metro de diámetro. Puesto que sólo el 10 por ciento de la luz que incide sobre el objetivo se absorbe y cumple su misión inutilizadora, el láser deberá arrojar 10.000 joule por centímetro cuadrado y por destello contra su blanco. El área cubierta por el haz luminoso a 1000 kilómetros de distancia vendría a coincidir con la superficie del espejo: unos 8000 centímetros cuadrados. Para concentrar un flujo de energía de 10.000 joule por centímetro cuadrado sobre este área, la energía total del haz luminoso tendría que rondar los 80 millones de joule por destello. Si los pulsos duraran unos 100 microsegundos, la potencia del láser sería de casi un millón de megawatt, cifra absolutamente inalcanzable. (Las grandes centrales nucleares de explotación comercial tienen una capacidad de generación de poco más de 1000 megawatt.)

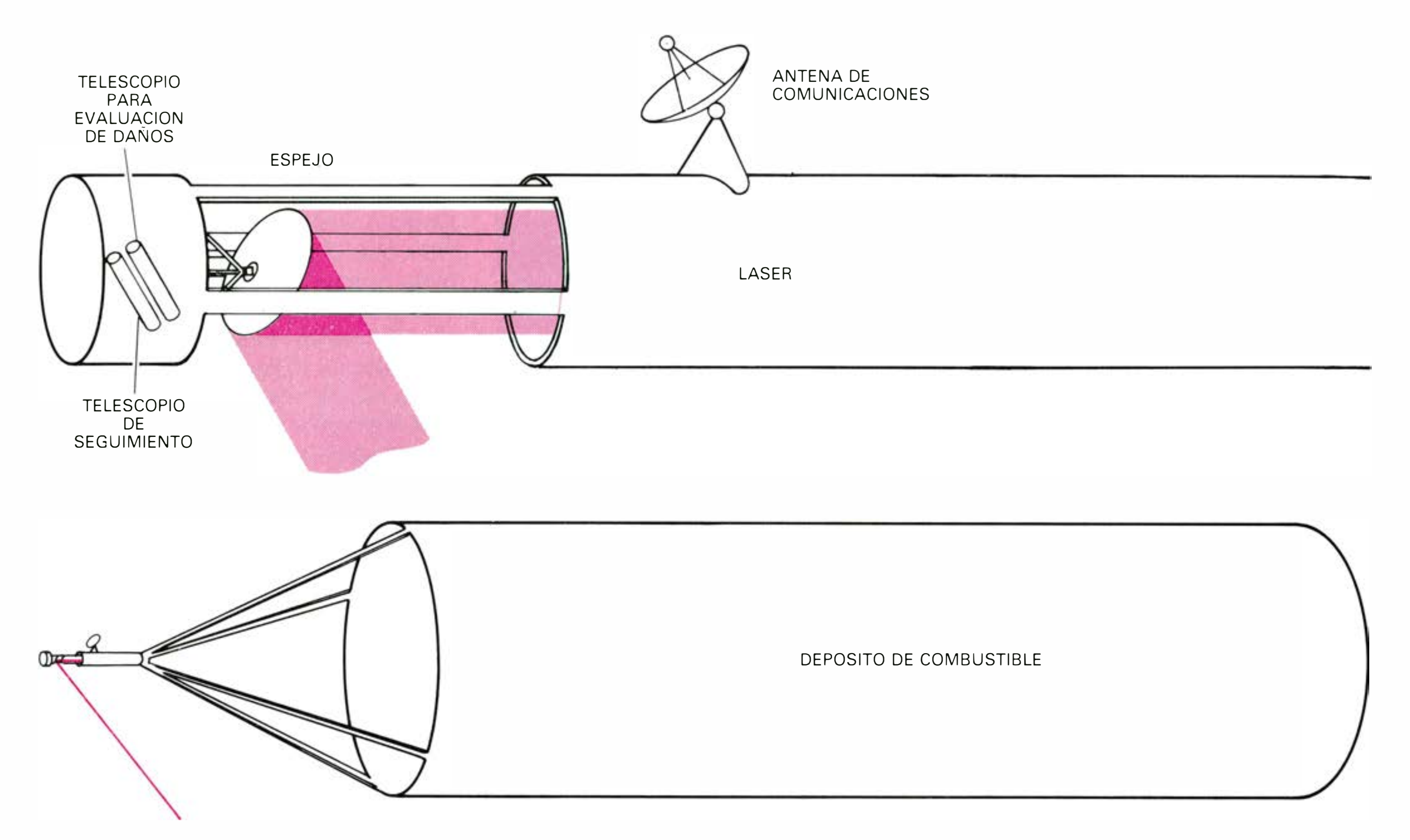
Si en vez de romper la carcasa del blanco la fundiéramos en una región circunscrita, no tendría que absorberse tanta energía aportada por el haz incesante. Así, una envoltura de aluminio de dos milímetros de espesor se fundiría cuando hubiese absorbido unos 400 joule por centímetro cuadrado. Si se supone que la reflectividad del aluminio es del 90 por ciento, un láser de dióxido de carbono de 100 megawatt necesitaría aproximadamente 100 segundos para asestar ese golpe a un objetivo distante 1000 kilómetros. La proporción del rendimiento es claramente desfavorable, puesto que el arma láser dispone de sólo medio segundo como máximo para atacar a cada misil en fase de despegue.

Podrían sortearse esas dificultades agrandando el espejo de puntería. Con un espejo de cuatro metros de diámetro, un láser de fluoruro de hidrógeno de 100 megawatt podría fundir el blanco en un segundo aproximadamente. Pero la fabricación de un espejo así, suficientemente robusto y de la necesaria calidad óptica, trasciende las posibilidades técnicas de los Estados Unidos y de cualquier otra nación. No se ve en el horizonte ninguna perspectiva de que se fabrique un espejo de cuatro metros dotado de la calidad óptica precisa.

Las necesidades de combustible de un sistema de arma láser representan otro obstáculo insuperable. Aun cuando el láser y su sistema de producción de energía funcionaran con absoluta eficacia, un láser de onda continua de fluoruro de hidrógeno consumiría algo así como 660 kilos de combustible por cada misil destruido. Para derribar 1000 misiles, habría que aprovisionar, a cada satélite, con 660 toneladas de combustible, que representan unas 20 cargas de la lanzadera espacial norteamericana. Los 50 satélites necesarios para garantizar una vigilancia continua de las bases soviéticas requerirían 1000 vuelos de la lanzadera espacial sólo pa-



LASER DE DESCARGA DE ELECTRONES. En este caso se logra la inversión de la población en las moléculas en el gas activo por medio de un haz de electrones. Los electrones ceden energía a las moléculas a través de las colisiones, llevándolas a un estado excitado. Se provoca la emisión de electrones mediante filamentos calientes o descarga eléctrica; un par de electrodos aceleran los electrones en la cavidad.



SATELITE ARTILLADO con un arma láser; necesitaría varios componentes, amén del propio láser. Así, un telescopio óptico o de infrarrojos para detectar y seguir la trayectoria del misil durante la etapa propulsada de su vuelo, y un espejo sobre articulación de rótula para apuntar el haz sobre el blanco. Tendría que tratarse de un espejo grande, robusto, altamente reflectante y ópticamente perfecto. Un sistema de control recibiría las señales del telescopio de seguimiento y dirigiría el haz moviendo el espejo. Otros sistemas de telescopios y captadores medirían el error de puntería o evaluarían el daño causado al blanco enemigo. Un enlace de comunicaciones con tierra transmitiría información acerca del objetivo y recogería las órdenes. Con mucho, empero, el componente más voluminoso del arma láser sería el relativo a la provisión de combustible y el sistema para manipularlo y suministrarlo al láser.

ra sus depósitos de combustible. La flotilla de cuatro lanzaderas espaciales, que realizaran cada una dos viajes por año, tardaría 125 años en suministrar el combustible necesario.

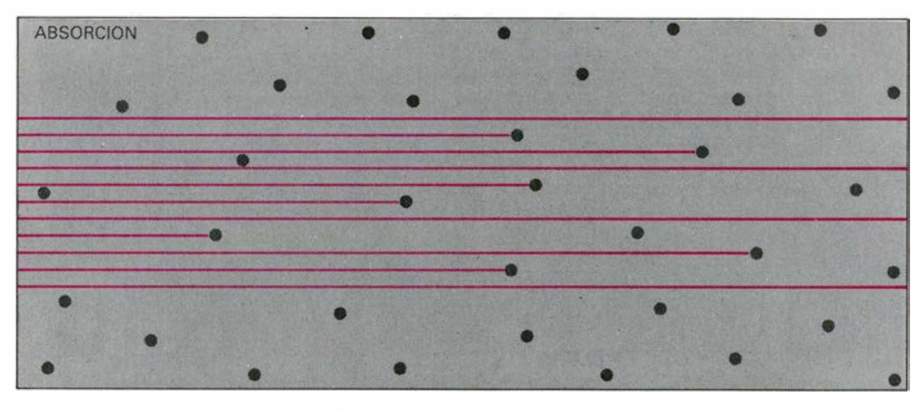
Las suposiciones en que se funda este razonamiento de un sistema hipotético de defensa contra misiles son ilusoriamente optimistas. Habría que puntualizar, en primer lugar, que no existe un láser de fluoruro de hidrógeno de 100 megawatt, ni hay visos de que un ingenio tal pudiera desarrollarse en un futuro previsible. Más aún, el rendimiento de un láser y del sistema de aporte energético no se aproximarán nunca al 100 por ciento. La eficiencia de los láseres existentes se halla en un porcentaje muy bajo; cabría esperar que algún día se alcanzara el 30 o 40 por ciento. A lo máximo a que puede aspirar un sistema de aporte energético es a un rendimiento del 30 por ciento. Así pues, el almacenamiento total de energía para cada satélite habría de multiplicarse por un factor de 10, como mínimo, y, más probablemente, de 30.

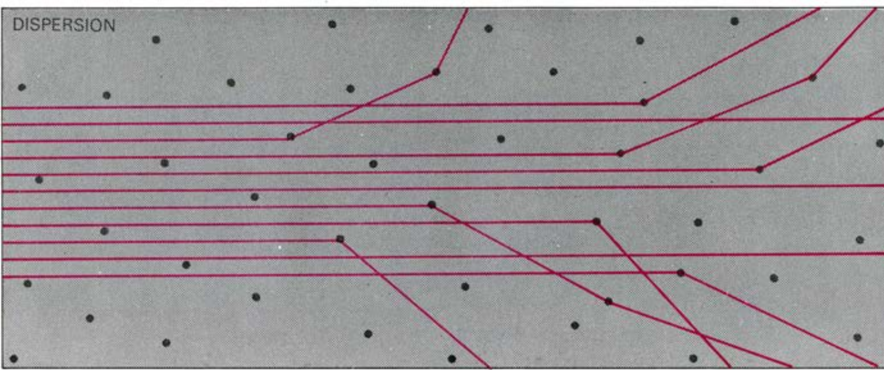
No es descabellado pensar que, andando el tiempo, se construya un arma láser apta para su despliegue en el espacio. Aun así, dudo que tenga éxito su explotación: sería vulnerable a múltiples contramedidas relativamente simples y baratas. Durante el largo tiempo invertido en ensamblar cada plataforma en el espacio, el sistema constituiría un fácil blanco a los ataques por armas antisatélites que hicieran explosión en su proximidad. Hasta una red terminada podría quedar transitoriamente inutilizada en instantes cruciales. ¿Cómo? Cegando sus captadores, saturando sus comunicaciones o confundiendo a sus sistemas de detección y seguimiento.

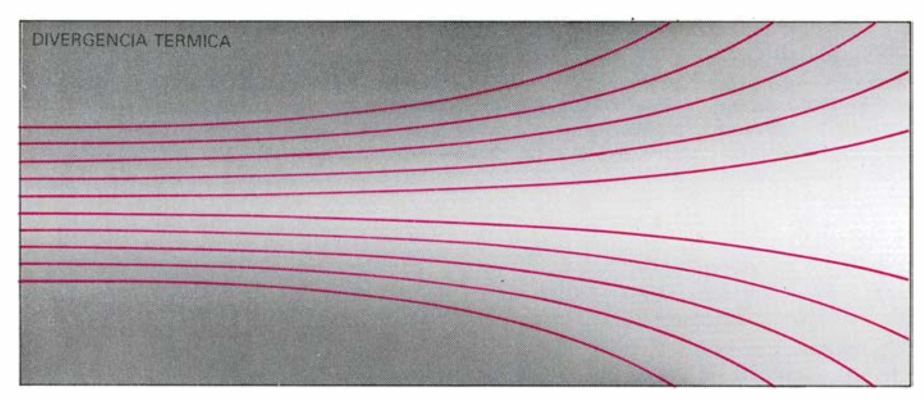
Otro destino verosímil para un arma láser con base en el espacio es el sistema antisatélite. Aunque su viabilidad resulta altamente discutible. En primer lugar, los satélites en órbita son ya vulnerables a las armas con cabeza nuclear, que pueden colocarse con exactitud en el espacio o incluso autodirigirse sobre un objetivo caliente en órbita. No se escaparía de ese riesgo el láser contra satélites con base en el espacio. El sistema láser sería, además, complejo y frágil y, por consiguiente, caro y difícil de mantener. Resulta muy improbable que un láser contra satélites ostente nunca una relación eficacia/coste mayor que ingenios mecánicos contra satélites lanzados desde tierra.

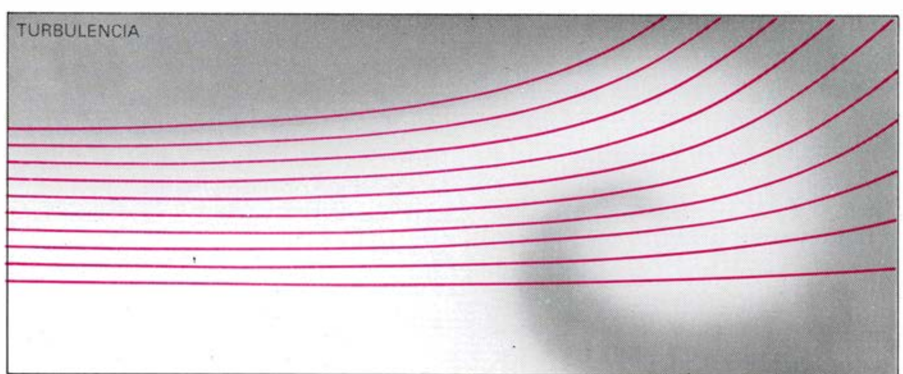
Volvamos a la viabilidad de un láser de alta energía que cumpliera misiones bélicas en la atmósfera terrestre. En principio, podría encargársele la protección de aviones y buques contra misiles enemigos y la destrucción de sus aeronaves. Pero ya hay sistemas de armas que realizan la misma tarea, incluidos los misiles supersónicos con guía de precisión y cañones de tiro rápido. Se trata de averiguar, por tanto, si el láser sería mejor arma. ¿Podría proporcionar esa protección de corto alcance con un coste menor y un rendimiento mayor?

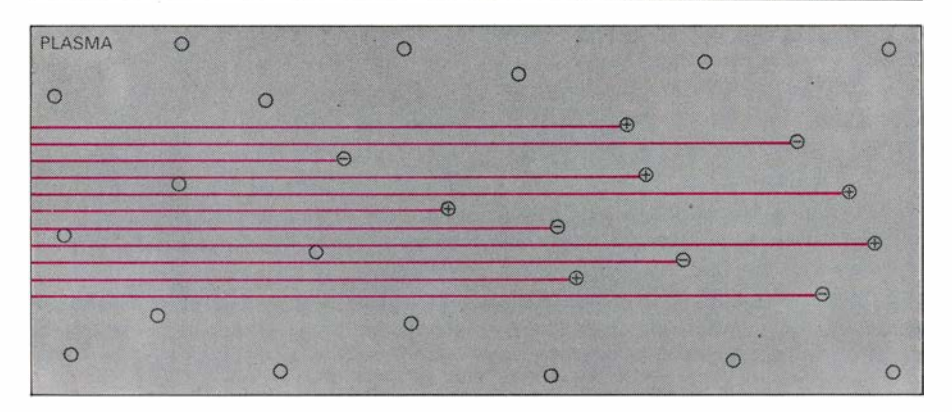
Lo primero que hay que poner en claro es la física de la propagación de un haz láser en la atmósfera. He descrito ya la divergencia térmica, la absorción y la ionización atmosférica. Sólo la divergencia y la absorción reducirían ya la intensidad de un rayo láser infrarojo a una distancia de cinco kilómetros en un factor de 100 a 300. Si el tamaño del haz en el objetivo es aproximadamente igual al tamaño del espejo transmisor, la intensidad del haz en el espejo habría de ser de 100 a 300 veces mayor de lo que se precisa que sea en el blanco. De esta necesidad surge otra: si









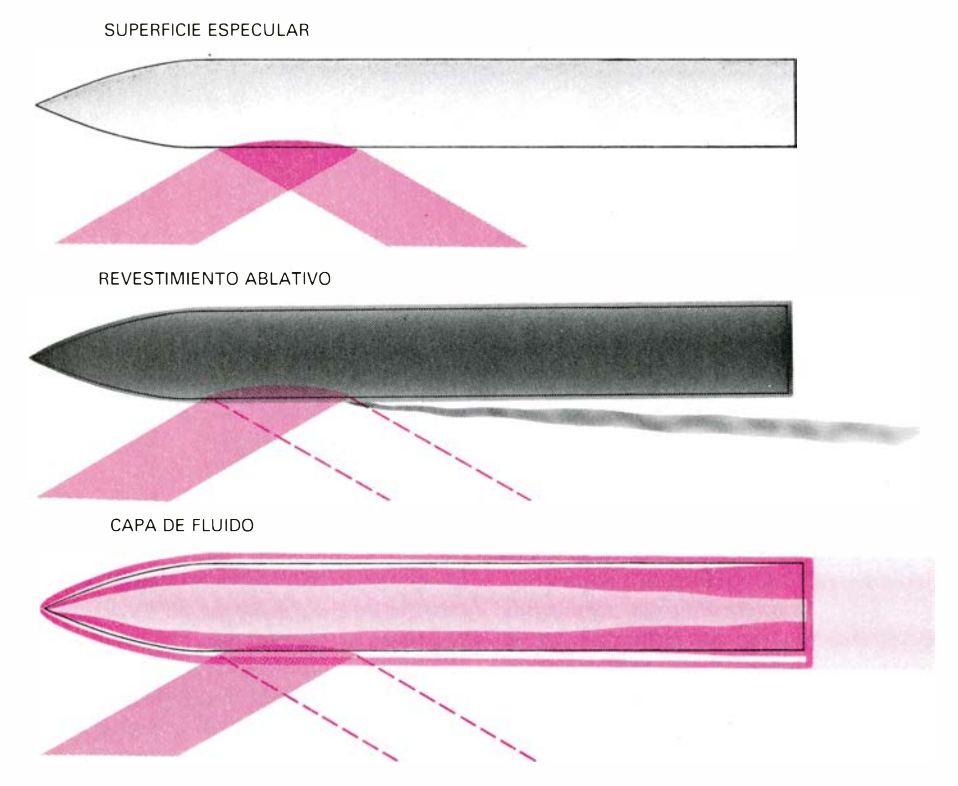


la luz láser ha de destruir el objetivo sin causar daño al espejo, la reflectividad de éste tendrá que ser de 200 a 300 veces mayor que la del objetivo. Ahora bien, imaginemos que el contrario protege sus misiles y aviones con un revestimiento reflectante. ¿Qué ocurrirá? Pues que el objetivo vendría a brillar tanto como el propio espejo.

Se podría intentar superar esta dificultad construyendo un gran espejo que enfocara el haz sobre el objetivo a pocos kilómetros de distancia. La sección transversal del haz en el blanco sería, así, menor que la misma sección en el espejo. Se trata, empero, de una solución desaconsejable, porque nada hay tan engorroso como manejar y apuntar espejos de más de un metro de diámetro, por no citar su vulnerabilidad a proyectiles simples. Más aún, los espejos con gran distancia focal pierden eficacia: la turbulencia atmosférica perturbaría y desenfocaría un haz luminoso ancho. Cualquier esfuerzo por construir espejos más reflectantes y para refrigerarlos podrá contrarrestarse pintando los misiles objetivo con varias capas de algún material ablativo que se quemaría y dispersaría la mayor parte de la energía incidente del haz láser. En esta competición entre láseres perfeccionados y contramedidas, aquéllos están en desventaja intrínseca, puesto que incluso con buen tiempo la atmósfera trabaja contra ellos.

Otra táctica a emplear con un láser comprometido en la defensa contra misiles sería esperar que el objetivo se hallara a sólo un kilómetro aproximadamente, antes de atacarlo. La intensidad del haz luminoso se degradaría entonces por un factor del orden de 10, en vez de 300, y quizá pudiera un espejo bastante reflector soportar el flujo de energía requerido para destruir el objetivo. El talón de Aquiles de este plan es la falta de tiempo. Un misil que se aproxime, por ejemplo, a dos veces la velocidad del sonido recorre el último kilómetro de su vuelo en segundo y medio. El láser carecería de tiempo su-

LAS INTERFERENCIAS atmosféricas podrían desviar o reducir la intensidad de un haz luminoso láser, que se propagase en su seno. El haz sería absorbido por partículas de materia y dispersado por el polvo, humo y gotículas de agua; las propias moléculas del aire absorben y dispersan al haz. La divergencia térmica se origina por el calentamiento del aire en el haz luminoso, creando una región de baja densidad, determinante de ese fenómeno de divergencia. La turbulencia conduce a variaciones locales en densidad y en índice de refracción que desvían y difunden el haz de forma impredecible. Una luz láser muy intenso podría ionizar las moléculas del aire creando un plasma que absorbería y, por consiguiente, interrumpiría la acción del haz.



MEDIDAS DEFENSIVAS, muy baratas, para proteger a los misiles de los efectos de la luz láser. Un método obvio: dotar al blanco potencial de una superficie altamente reflectante, para así absorber sólo una fracción muy pequeña de la luz. También: recubrir el misil de un revestimiento ablativo, que se quemaría y arrastraría lejos la energía de la radiación láser. Una capa de fluido secretada desde el extremo de la cabeza produciría el mismo efecto. Otra estratagema: que el misil gire alrededor de su eje.

ficiente para enfrentarse a más de un misil atacante. En el mismo intervalo, un cañón de tiro rápido podría disparar varios proyectiles explosivos contra el blanco.

Aun cuando la luz láser viaja casi un millón de veces más deprisa que un proyectil ordinario, el arma láser no gozaría de ninguna ventaja operacional intrínseca sobre un cañón de tiro rápido para la protección contra misiles a corta distancia. Tiene el láser, por contra, varias desventajas. Un misil atacante puede ir protegido contra la luz láser (particularmente de un haz luminoso continuo de baja intensidad) mediante una fina película de alguna sustancia que se excrete continuamente en la punta del misil, para absorber la energía del haz luminoso y disipar el calor. El misil podría también girar de suerte que dispersara el calor por toda su superficie. Por último, hasta para una distancia al objetivo del orden de un kilómetro el mal tiempo puede neutralizar completamente un arma láser.

Como postrera consideración pensemos que el proceso de detección y seguimiento de un objetivo exige más de un láser que de otra arma defensiva, ya que el haz debe incidir realmente sobre el objetivo para que resulte eficaz. La exactitud convencional del sistema de seguimiento de un cañón que dispara proyectiles explosivos es mucho menos exigente, en particular si el proyectil lleva un rastreador de infrarrojos que le capacite para autodirigirse sobre el objetivo.

De la comparación se desprende que las armás láser que operan en la atmósfera no ofrecen claras ventajas sobre las ya existentes para la defensa a corta distancia. Puede bloquear su acción el mal tiempo, no operan con eficacia más allá del alcance de unos pocos kilómetros, se neutralizan por las contramedidas con mayor facilidad que los proyectiles ordinarios o misiles supersónicos y requieren un sistema de seguimiento mucho mejor. En esas condiciones, cuesta imaginar que el desarrollo y el despliegue de armas tan frágiles, complejas y caras fortalecerían la potencia militar de una nación.

Pero sería falso deducir de lo expuesto que la investigación sobre los láseres de alta energía carezca de sentido. Que no se les vea un futuro prometedor en el campo de batalla, no empece que posean aplicaciones útiles en la industria, particularmente en ingeniería química y en sistemas energéticos basados en la fusión nuclear. Por estas razones, y no por aplicaciones militares quiméricas, los Estados Unidos harían bien en continuar ahondando en muchos aspectos de la tecnología de los láseres de alta energía.

Navegación magnética en las bacterias

Ciertas bacterias acuáticas son magnetotácticas: poseen pequeñas brújulas internas que las orientan en el campo magnético terrestre. El fenómeno quizá pueda extrapolarse a organismos superiores

Richard P. Blakemore y Richard B. Frankel

Las bacterias se cuentan entre los organismos más sencillos. Lo que no ha sido óbice para que desplegaran una amplia gama de adaptaciones sensoriales y motoras, denominadas respuestas tácticas, por las que interaccionan con su entorno. Muchas bacterias son quimiotácticas: tienden a desplazarse hacia una concentración superior o inferior de una determinada sustancia. Otras presentan fototaxia: se acumulan en una región iluminada.

A lo largo de los últimos años se ha sometido a investigación una nueva respuesta táctica, hoy confirmada: hay bacterias acuáticas que son magnetotácticas; dicho de otro modo, tienden a nadar a lo largo de líneas de campos magnéticos. Se hallan sujetas, por tanto, al influjo del campo magnético terrestre. Se ha demostrado que las bacterias magnetotácticas sintetizan, para luego acarrearlos en su interior, pequeños cristales de magnetita (Fe_3O_4), el óxido de hierro conocido por los primeros navegantes como piedra imán. Se han encontrado también partículas de magnetita en otros organismos. Por tanto, el descubrimiento de magnetotaxia en las bacterias da pie a la sospecha de que el hombre no es el único organismo que ha dominado el arte de la navegación magnética.

Más de mil años antes de la invención de la brújula magnética en la Edad Media, chinos y griegos de la antigüedad conocían ya las propiedades magnéticas de la magnetita. El material en que se basaban las primeras brújulas era una pieza alargada de magnetita con una gran inmantación natural. Se descubrió luego que una aguja de hierro magnetizada por frotamiento con magnetita podía servir para los mismos fines.

La idea de que la Tierra misma constituía un imán no llegaría hasta 1600, cuando William Gilbert la propuso en *De magnete magneticisque corporibus et de magno magnete tellure* (*Sobre el imán y los cuerpos magnéticos y sobre el gran imán de la Tierra*). Gilbert, médico de la reina Isabel I de Inglaterra, basó su tratado en 16 años de estudio sobre las interacciones entre agujas magnetizadas y esferas hechas de magnetita. Denominó a esas esferas "teleras", pequeñas tierras.

Algunas décadas antes, Robert Norman, un fabricante de brújulas inglés, había observado que si a una aguja magnetizada se la permitía girar libremente sobre un eje en el plano vertical norte-sur, el extremo de la aguja que buscaba el norte se inclinaba por debajo del horizonte. En Londres, por ejemplo, el ángulo de inclinación era aproximadamente de 70 grados. Gilbert explicó la inclinación como un efecto de la geometría del magnetismo de la Tierra. Exploró dicha geometría midiendo la orientación de agujas inmantadas colocadas cerca de la superficie de una telera. ¿Qué halló? Pues que el campo magnético inducido por una esfera magnetizada era radial (vertical) en los polos, tangencial (horizontal) en el ecuador y, entre ambos extremos, inclinado un ángulo cuyo valor aumentaba con la latitud. (Sabemos hoy que los polos magnéticos de la Tierra no coinciden exactamente con los polos geográficos, discrepancia que no afecta a nuestro trabajo, por cuya razón la pasaremos por alto.)

Se ha convenido en que el sentido de un campo magnético sea el que señale el extremo que busca el norte de la aguja magnética de una brújula. Por tanto, el campo geomagnético apuntará hacia arriba en el polo Sur, hacia abajo en el polo Norte y horizontalmente hacia el norte en el Ecuador. En el hemisferio Norte, el campo estará inclinado hacia abajo, y, hacia arriba, en el hemisferio Sur.

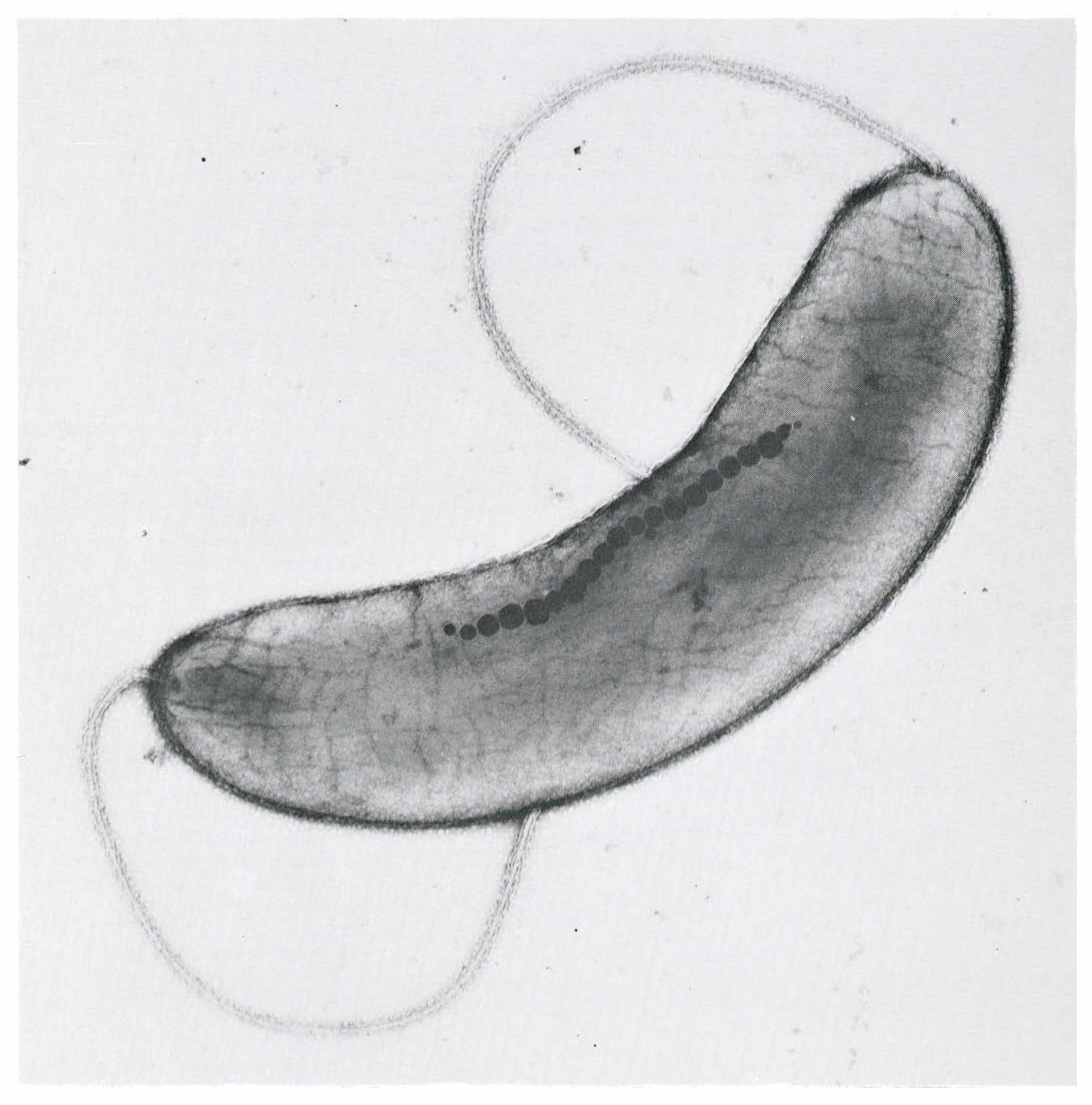
La primera indicación de que las bacterias eran sensibles al campo geomagnético la advirtió, en 1975, Richard P. Blakemore, a la sazón alumno de segundo ciclo del departamento de microbiología que la Universidad de Massachusetts tiene en el campus de Amherst. (Blakemore es uno de los firmantes de este artículo.) Se hallaba estudiando las bacterias que suelen colonizar el cieno de pantanos salobres, cuando observó que algunos microorganismos nadaban persistentemente en un mismo sentido, a través del campo visual del microscopio, para concentrarse en el borde de una gota de agua fangosa. Que no se trataba de una respuesta fototáctica lo evidenciaban las bacterias en su avance hacia el mismo borde sin tener en cuenta la distribución de la luz que llegaba a la preparación microscópica. Podíamos cubrir el microscopio con una caja, darle la vuelta o transportarlo a otra habitación, las bacterias seguían nadando en la misma dirección geográfica. Todo indicaba que estaban dirigidas por algo ubicuo, posiblemente el campo magnético de la Tierra.

Que el comportamiento de tales bacterias era magnetotáctico resultó fácil demostrarlo. A tal fin examinamos una gota de agua con barro al microscopio con poco aumento (unos 80 diámetros), con iluminación sobre campo oscuro. Las bacterias móviles, que refractaban la luz, aparecían, bajo esas condiciones, cual puntos luminosos en movimiento. En ausencia de todo campo magnético que no fuera el geomagnético, algunas de las bacterias nadaban persistentemente en dirección norte y se acumulaban en el borde norte de la gota. Si se aproximaba una barra

magnética, las bacterias nadaban hacia el polo que atrae el extremo de una aguja inmantada que señala al norte. Y a la inversa, nadaban apartándose del polo que atrae el extremo de la aguja que busca el sur.

Al principio, utilizábamos pequeños imanes permanentes para observar la respuesta de los organismos a los campos magnéticos y separarlos de los sedimentos para su posterior análisis. Más tarde, en el Instituto Oceanográfico de Woods Hole, Adrianus J. Kalmijn, especializado en biofísica sensorial de animales marinos, se sumó al proyecto diseñado para explorar la conducta magnetotáctica de las bacterias en campos magnéticos aproximadamente equivalentes en intensidad al campo geomagnético. (El campo magnético de la Tierra tiene una intensidad del orden de un gauss.) En estos experimentos se inducían campos magnéticos uniformes mediante un par de bobinas de Helmholtz, que permitían controlar el sentido y la intensidad de los campos [*véase la ilustración de la página siguiente*]

Cuando se alinea el eje de las bobinas de Helmholtz en la dirección norte-sur, la intensidad del campo horizontal que actúa sobre las bacterias es la suma del campo generado por las bobinas más la componente horizontal del campo geomagnético. Si el campo genera-



BACTERIA MAGNETOTACTICA, según una micrografía electrónica de transmisión que revela una notable característica de su estructura interna: una cadena de partículas opacas al haz de electrones (*manchas negras*). Se trata de los magnetosomas, partículas formadas por el óxido de hierro conocido como magnetita (Fe_3O_4). Puede apreciarse el paralelismo que guardan con el eje mayor de la célula. La cadena funciona como si fuese una brújula: orienta a la bacteria en la dirección de las líneas de fuerza del campo magnético terrestre. La especie que aquí se muestra tiene un flagelo (un propulsor filamentoso) en cada extremo y puede nadar tanto hacia delante como hacia atrás. (Los extremos de los dos flagelos quedan, en esta imagen tomada por Richard P. Blakemore y Nancy Blakemore, bajo la bacteria, motivo por el cual no los recoge la micrografía.) El aumento es de unos 56.000 diámetros.

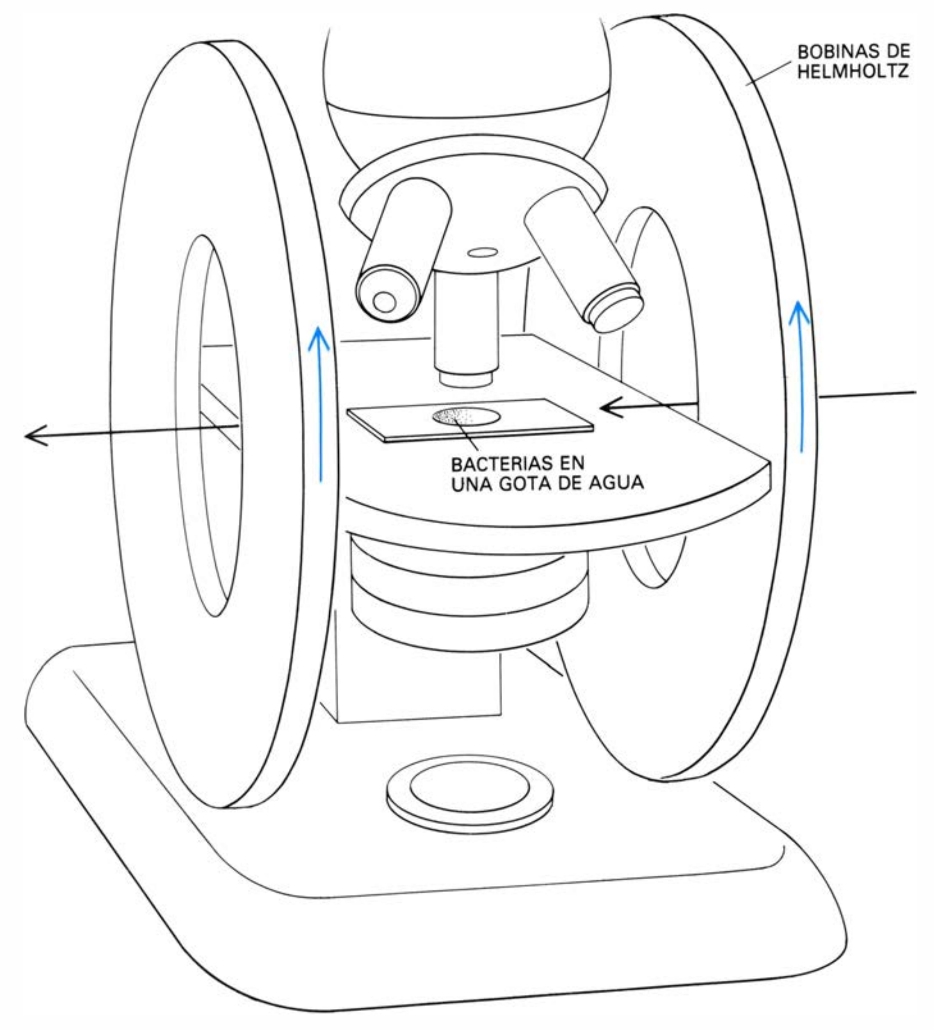
do por las bobinas es algo más intenso que el campo geomagnético, el sentido neto de los campos combinados coincidirá con el del campo aplicado. Las bacterias nadan a lo largo de líneas magnéticas de fuerza en el sentido del campo (es decir, en el sentido indicado por el extremo que busca el norte de una aguja de una brújula). Cuando se invierte la corriente eléctrica de las bobinas, se invierte el campo magnético; las bacterias dan entonces un giro de 180 grados y nadan en sentido contrario. Cada vez que se invierte el campo magnético, las bacterias dan una vuelta en forma de U, de manera que siempre están nadando en el sentido del campo. Puede afirmarse, por tanto, que los organismos buscan la dirección norte.

La persistente natación de las bacterias magnetotácticas a lo largo de las líneas del campo magnético contrasta con los movimientos en vaivén característicos de las bacterias quimiotácticas, de las que *Escherichia coli* constituye un ejemplo. Aun después de su muerte, las bacterias magnetotácticas adoptan una orientación constante dentro del campo magnético y giran 180 grados cuando se invierte la corriente de las bobinas (las bacterias muertas, sin embargo, no emigran hacia el norte ni hacia el sur).

En pocas palabras, las bacterias actúan a modo de pequeños dipolos magnéticos biológicos.

En sedimentos tomados de pantanos, cenagales, plantas de tratamiento de aguas residuales o lugares similares, la concentración de bacterias magnetotácticas se cifra usualmente entre 100 y 1000 células por mililitro. Cuando los sedimentos se almacenan en frascos de laboratorio, la concentración aumenta, en unas pocas semanas, de 100.000 a un millón de células por mililitro. Esta elevada densidad de población persiste con frecuencia hasta dos años sin que se añadan nutrientes. Puede aprovecharse la respuesta magnetotáctica de las bacterias para hacerlas nadar, apartándolas del sedimento y de otros desperdicios, e introducirlas en una gota de agua limpia, de donde pueden recogerse para su examen al microscopio.

Se han hallado más de una docena de tipos morfológicamente distintos en muestras así preparadas. Comprenden cocos (bacterias esféricas), bacilos (bacterias en forma de bastón) y espirilos (bacterias helicoidales). Esta observación, junto con el hecho de que las bacterias magnetotácticas se encuentran tanto en sedimentos de aguas dulces como de aguas marinas, indica que la magnetotaxia constituye una propiedad compartida por varias especies bacterianas.

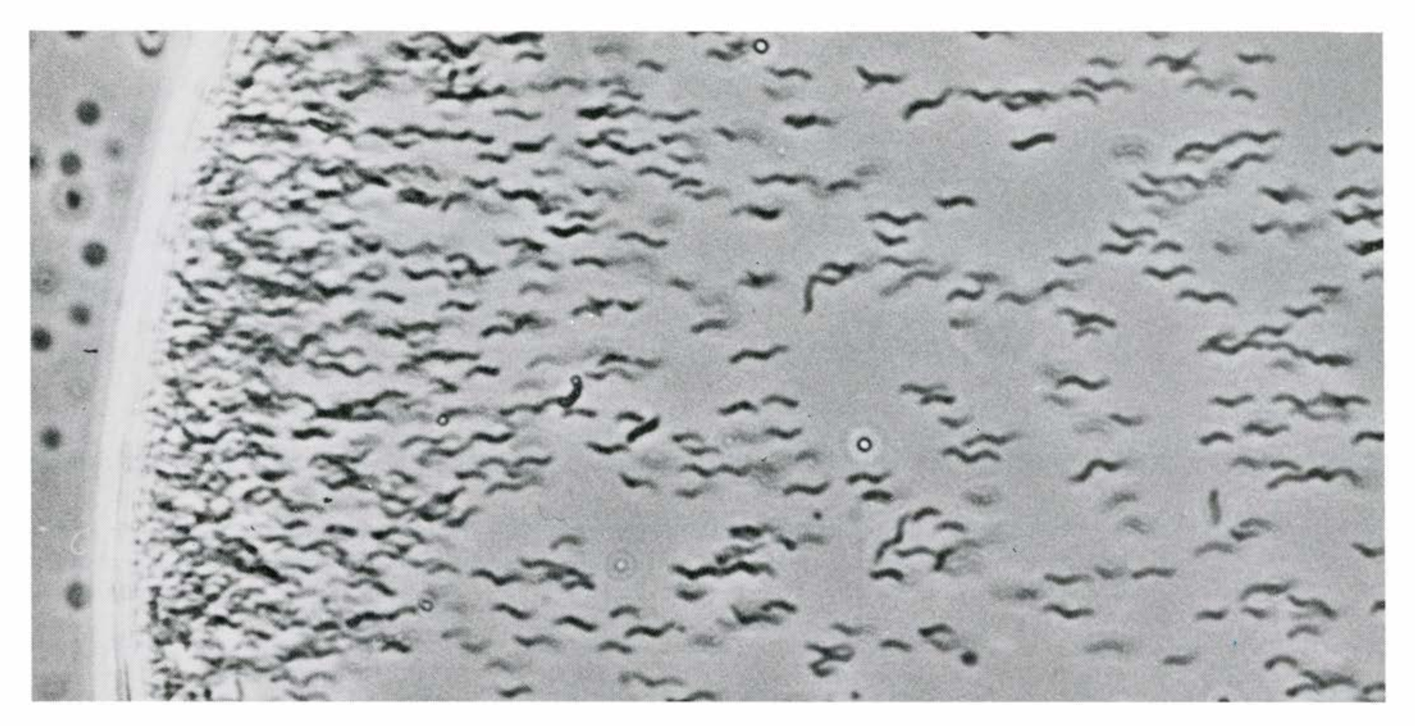


DISPOSITIVO EXPERIMENTAL empleado por el equipo de los autores para estudiar el comportamiento de las bacterias magnetotácticas en un campo magnético aplicado. Consta de un microscopio que lleva acoplado un par de bobinas de Helmholtz. La corriente eléctrica de las bobinas (*flechas de color*) crea un campo magnético uniforme a lo largo de su eje (*flecha negra*). Si se invierte la corriente, se invierte también el campo magnético. Variando la magnitud de la corriente se controla la intensidad del campo.

Uno de nosotros (Blakemore) y Ralph S. Wolfe, de la Universidad de Illinois, han aislado y cultivado en un medio artificial una de ellas. Se trata de un espirilo magnetotáctico de agua dulce, bautizado con el nombre de *Aquaspirillum magnetotacticum*. Las bacterias de esta especie poseen un flagelo (un propulsor filamentoso) en cada uno de sus extremos; pueden nadar tanto hacia delante como hacia atrás. Esta capacidad las distingue de otras muchas bacterias magnetotácticas, que tienen el aparato flagelar sólo en un extremo de la célula y parecen impulsadas siempre hacia delante.

Un examen más detenido, al microscopio electrónico de transmisión, de la estructura interna de la bacteria *A. magnetotacticum*, nos revela una característica muy notable: la presencia de aproximadamente 20 partículas, opacas al haz de electrones y de forma cúbica u octaédrica, dispuestas en una cadena paralela al eje mayor de la célula. En todas las bacterias magnetotácticas que se han examinado hasta ahora mediante esta técnica se aprecian tales partículas en el citoplasma de la célula. Las partículas suelen aposentarse en una sola cadena, o a lo sumo dos; muestran un destacado parecido en cuanto a forma y tamaño. Cada partícula está rodeada por una vaina o membrana, adyacente (contigua, a veces) a la membrana citoplasmática que envuelve toda la célula. Las partículas encerradas, que miden unos 500 angstrom de anchura, se conocen con el nombre de magnetosomas.

Micrografías electrónicas tomadas por David L. Balkwill y Denise Maratea, de la Universidad de New Hampshire, revelaron que, en el caso de *A. magnetotacticum*, la vaina del magnetosoma constaba de una fina capa opaca a los electrones, separada de la



BACTERIAS BUSCADORAS DEL NORTE nadan en la dirección de un campo magnético inducido y se acumulan en el borde de una gota de agua en una demostración de la magnetotaxia. De no mediar ese campo provocado, las bacterias nadarían persistentemente hacia el norte bajo la influencia del campo magnético terrestre. Se las ha bautizado con el nombre de *Aquaspirillum magnetotacticum*. Aparecen, en esta imagen, aumentadas unas 1700 veces.

propia partícula por un espacio transparente a los electrones. En otras especies de bacterias magnetotácticas se aprecian tres capas en la membrana que encierra a cada partícula. La composición química de la envoltura de los magnetosomas es aún objeto de estudio. Quizá se cuente, entre sus funciones, la de mantener cada partícula en una posición fija del interior celular.

¿Cómo determinar la composición elemental de las partículas magnetosómicas? Mediante la observación de los rayos X emitidos cuando, sobre una muestra, incide el haz de electrones del microscopio electrónico. Los electrones de alta energía colisionan con los electrones de los átomos de la muestra, desplazándolos. El reordenamiento de los electrones atómicos subsiguiente va acompañado de la emisión de rayos X, que tienen un espectro de energías característico del elemento del que emanan. Al estimular con electrones los magnetosomas de *A. magnetotacticum*, predominan los rayos X característicos del hierro. Cuando se exponen de igual manera regiones de la célula que carecen de magnetosomas, el espectro de rayos X emitido no incluye los característicos del hierro. De ello se desprende que las partículas deben estar compuestas de alguna forma de hierro. De hecho, este elemento supone el 2 por ciento del peso seco de la célula.

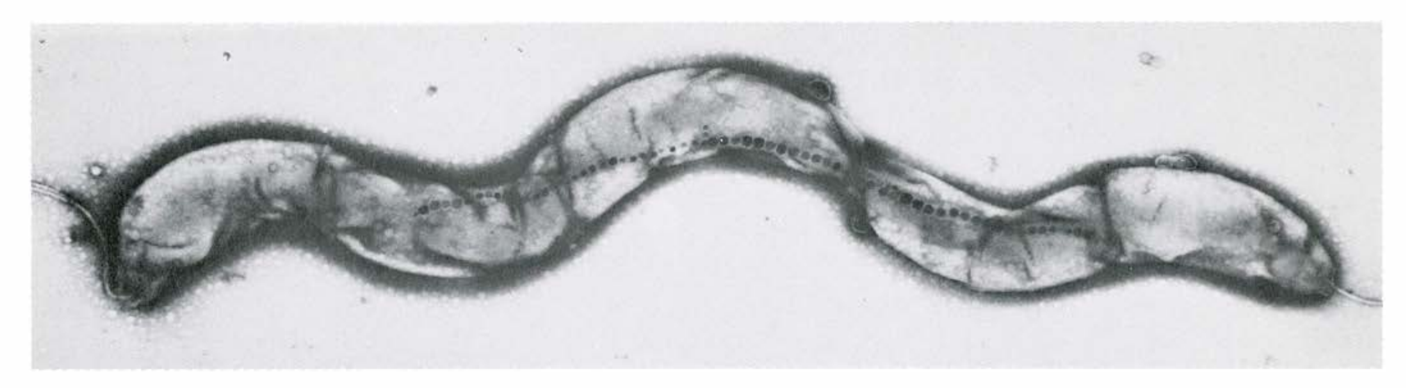
El experimento que detallamos a continuación confirma la conexión entre los magnetosomas y la respuesta magnetotáctica. Si se aporta al medio de cultivo de *A. magnetotacticum* hierro en forma de un complejo órganico soluble, a una concentración de uno o dos miligramos de hierro por litro, las bacterias crecen, sintetizan magnetosomas y son magnetotácticas. Si la misma cepa de bacterias se inocula en un medio de cultivo que no contenga más de medio gramo de hierro por litro, las bacterias crecen como antes, pero no forman entonces magnetosomas ni son magnetotácticas. La concentración total de hierro en las células magnetotácticas ronda sólo el 10 por ciento de la que tienen las células magnetotácticas. Como es obvio, únicamente las bacterias con magnetosomas presentan magnetotaxia. El hierro requerido para la síntesis de magnetosomas se obtiene sin dificultad a concentraciones de un miligramo por litro, aproximadamente, en muchos ambientes naturales donde viven las bacterias magnetotácticas.

La forma química del hierro de los magnetosomas se determina por espectroscopía de Mössbauer. Esta técnica permite detectar, en el espectro de los rayos gamma que sufren corrimiento Doppler absorbidos por un núcleo atómico, los sutiles desplazamientos causados por la configuración de los electrones que rodean al núcleo. Puesto que la configuración electrónica está determinada por el entorno químico del átomo, el espectro de Mössbauer nos desvela la forma química que un elemento tiene en una determinada muestra.

En este caso, la forma química del hierro se determinó midiendo el espectro de Mössbauer del isótopo Fe-57 en células magnetotácticas de *A. magnetotacticum* sometidas a criofractura y en magnetosomas aislados de las células. El espectro resultó ser casi idéntico al de una muestra de laboratorio de magnetita pura [*véase la ilustración de la página 22*]. Se puede afirmar por tanto que el hierro de *A. magnetotacticum* se encuentra principalmente en la forma Fe_3O_4. Como era de esperar, las líneas de los espectros correspondientes a la magnetita no aparecían en los espectros de Mössbauer correspondientes a células no magnetotácticas.

Además de las líneas correspondientes a la magnetita, en el espectro de Mössbauer de las células magnetotácticas se dibujaban un par de líneas en el centro del espectro que correspondían a otro compuesto de hierro, presente, según se ve, en cantidades importantes. Puesto que el hierro que se aportó al medio de cultivo estaba en forma de complejo órganico soluble, la magnetita de las partículas era, sin la menor duda, producto de síntesis bacteriana. Para nosotros, el material responsable de las líneas adicionales en el espectro de Mössbauer ha de ser un precursor, que lleva hierro, de la ruta bioquímica por la que las bacterias sintetizan la magnetita.

Saber que las partículas magnetosó-



BACTERIA HELICOIDAL de la especie *A. magnetotacticum*, ampliada unas 25.000 veces en esta micrografía electrónica de transmisión. Esta especie magnetotáctica de agua dulce se ha podido aislar y cultivar en el laboratorio. El ejemplar que se muestra tiene una cadena excepcionalmente larga de magnetosomas dispuestos en fila a lo largo de su eje. Probablemente se halle en proceso de división celular, como apuntan los esbozos de septos transversales.

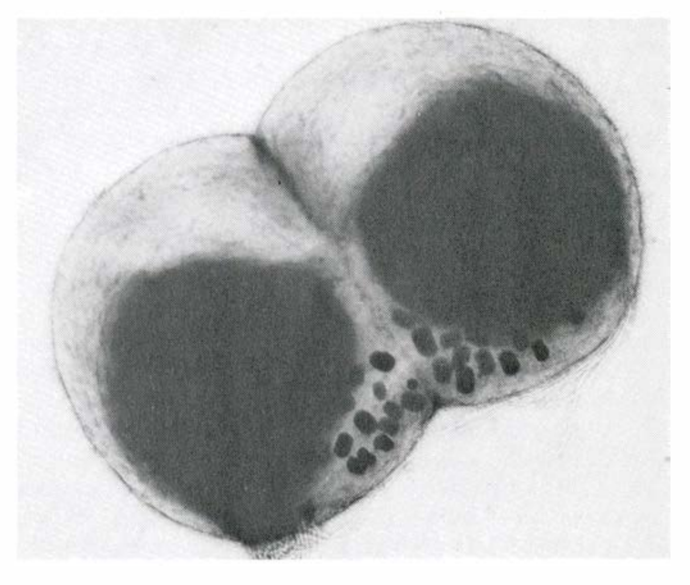
micas están formadas por magnetita nos permite aclarar el mecanismo de la magnetotaxia. La forma de magnetismo observada en la magnetita se denomina ferromagnetismo. Como en todos los materiales magnéticos, la fuente última del magnetismo es el momento magnético asociado al momento angular del espín de cada átomo de hierro. Las interacciones entre átomos originan la alineación paralela de los espines en una región del material. Las propiedades magnéticas macroscópicas de una partícula de magnetita dependen de su tamaño y de su forma. Si es grande, incluirá varios dominios magnéticos: regiones donde los ejes de espín están orientados de forma antiparalela. En un material ferromagnético está favorecida la configuración antiparalela, porque rebaja la energía magnética total de la partícula; reduce también el momento dipolar magnético residual.

Los dominios magnéticos están separados por regiones de transición denominadas paredes de los dominios. La anchura de tal pared, que viene determinada por ciertas propiedades magnéticas intrínsecas del material, no depende del tamaño de la partícula. Si las dimensiones de la partícula son menores que la anchura de la pared, no cabe la formación de dominios distintos y la partícula constituye así un solo dominio magnético. En tal partícula, el momento magnético residual no cambia con el tiempo y se acerca al máximo teórico para el material.

Una partícula de un tamaño todavía menor formará también un único dominio, pero su momento magnético eficaz podrá menguar por mor de otro fenómeno: el de superparamagnetismo. Son varias las orientaciones del momento magnético que pueden verse favorecidas por la estructura cristalográfica del material, dotadas todas ellas de la misma energía. En una partícula muy pequeña, el momento magnético fluctúa entre las orientaciones por causa de las disrupciones térmicas. En otras palabras, el momento magnético residual no es permanente en una partícula así. A una temperatura dada, el tiempo requerido para que el momento magnético se oriente es una función exponencial del volumen de la partícula. Si ésta posee un tamaño suficientemente grande, el tiempo de reorientación térmica viene a resultar infinito.

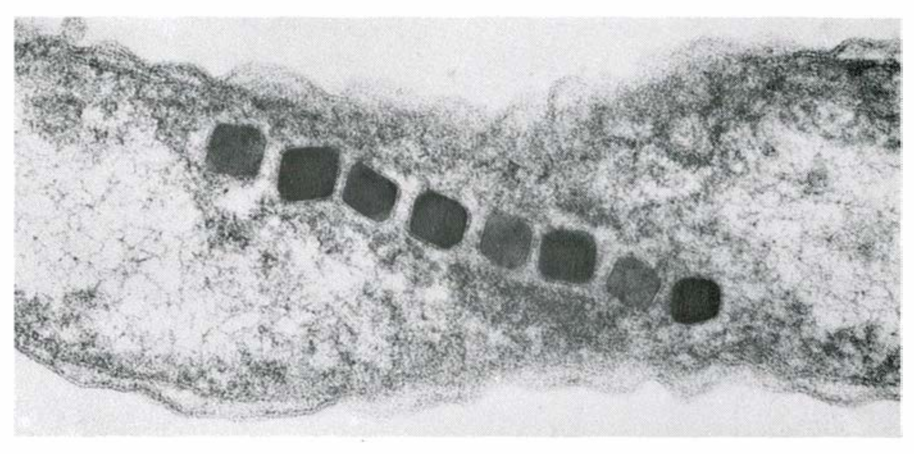
¿Qué cabe deducir de lo hasta aquí expuesto? Pues que una partícula con un solo dominio, y que ocupa un volumen demasiado grande para ser superparamagnética a la temperatura ambiente, posee un momento dipolar magnético que corresponde al máximo de magnetización para esa sustancia. En





BACTERIAS ESFERICAS, o cocos. Gozan también de magnetotaxia. Su citoplasma contiene partículas de magnetita. En algunos cocos, los magnetosomas forman cadenas (*izquierda*); en otros se apiñan en racimos (*derecha*). Ambas especies tienen haces apareados de flagelos a un lado y nadan solamente en un sentido. Los cuerpos grandes y opacos a los electrones que se aprecian en los cocos de la derecha poseen una alta concentración de fósforo y de potasio; se desconce su significación en la vida del organismo. También esta célula está a punto de dividirse. Micrografías de Blakemore y W. O'Brien.

el caso de la magnetita, la máxima dimensión superparamagnética se sitúa entre 300 y 400 angstrom, en tanto que la dimensión mínima para la formación de dominios múltiples está entre 800 y 1200 angstrom. Las partículas producidas por *A. paramagneticum* y otras especies de bacterias magnetotácticas caen precisamente dentro de la escala de tamaños dentro de la cual la magnetita se presenta como partícula termoestable de un solo dominio: entre los 400 y los 1200 angstrom. En el laboratorio es difícil obtener partículas de magnetita del tamaño de la gama de un solo dominio. Ni tampoco está claro cómo lo realiza la célula bacteriana. Suponemos que el tamaño máximo de la partícula deberá estar condicionado, en cierta manera, por la envoltura del magnetosoma y que, por consiguiente, se halla bajo control biológico.



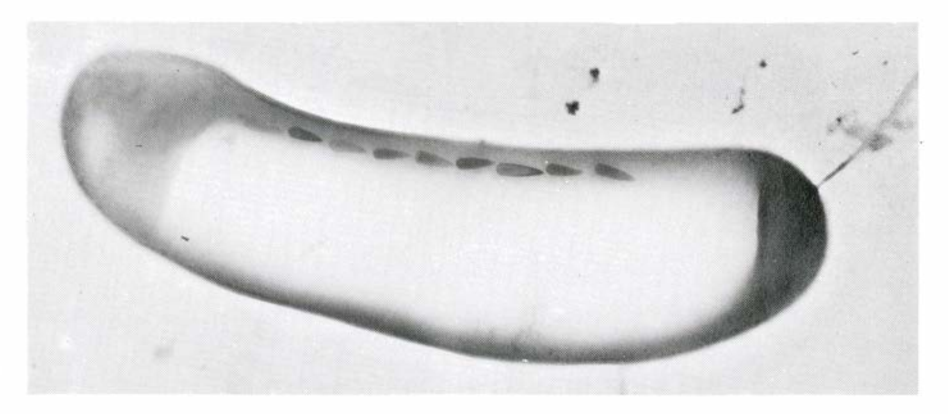
LA FINA MEMBRANA que rodea a cada partícula de magnetita apenas se distingue en esta micrografía electrónica de transmisión, muy ampliada, de una sección de bacteria magnetotáctica. Una función de la envoltura podría ser la de mantener a cada partícula en una posición invariable en el interior celular.

¿Qué decir acerca de la configuración de las partículas de magnetita que hay en el interior celular? Muchas especies de bacterias magnetotácticas, lo avanzamos antes, tienen una o dos cadenas de magnetosomas. Las interacciones magnéticas entre las partículas de una cadena tienden a orientar sus momentos dipolares magnéticos en paralelo a lo largo del eje de la cadena. Eso observaron Charles P. Bean e Israel S. Jacobs, del General Electric Research and Development Center, quienes, en un contexto diferente, consideraron las propiedades de una cadena de imanes de un solo dominio dos décadas antes del descubrimiento de las bacterias magnetotácticas. Debido a la orientación paralela de los momentos, el momento total de la bacteria es la suma de los momentos de las distintas partículas. Puesto que se conoce el momento por unidad de volumen de magnetita, podemos calcular el momento total por bacteria a partir del tamaño y del número de sus magnetosomas; éstos se aprecian con suficiente nitidez en las micrografías electrónicas.

La orientación de una bacteria magnetotáctica en agua a temperatura ambiente está determinada por la resultante del par de torsión, ejercido sobre el momento dipolar magnético de la bacteria por el campo geomagnético, y las fuerzas aleatorias que tienen su origen en el movimiento térmico de las moléculas de agua (el llamado movimiento browniano). El grado medio de orientación de un conjunto de dipolos magnéticos en libre rotación que se hallan en contacto térmico con su entorno fue calculado, a finales del siglo pasado, por el físico francés Paul Langevin. Demostró que, a medida que aumentaba la proporción de energía magnética respecto de energía térmica, crecía asimismo el valor promedio de la orientación del conjunto; para proporciones mayores de 10, el conjunto se aproximaba a la orientación completa.

En *A. magnetotacticum*, la proporción ronda en torno a 16, lo que significa una orientación media que raya el valor máximo. Además, el incremento del número de partículas por bacteria sólo aumentaría marginalmente el grado de orientación de las bacterias a lo largo de la dirección del campo magnético. Pero si las bacterias contuviesen sólo magnetita suficiente para que su concentración total de hierro fuera comparable con la de otras bacterias comunes, *E. coli* por ejemplo, resultaría muy difícil distinguir su orientación de una distribución al azar. Así pues, las partículas de un solo dominio dispuestas en una cadena constituyen una brújula biomagnética de suficiente potencia para que la bacteria, en un medio acuático a temperatura ambiente, se oriente de un modo eficaz en la dirección del campo geomagnético.

La explicación más sencilla del mecanismo de magnetotaxia se vale del siguiente razonamiento: la bacteria se deja arrastrar, en su movimiento de natación, por el par torsor ejercido sobre su brújula biomagnética por el campo magnético terrestre. Hipótesis que han confirmado mediciones de la velocidad de las células migrantes en función de la fuerza del campo magnético. Si la intensidad del campo era elevada (de varios gauss), las bacterias seguían fielmente el sentido del campo, a una velocidad de migración, en tal sentido, bastante alta. Cuando se trataba de un campo débil, la agitación térmica repercutía notablemente sobre las bacterias, su natación se tornaba más errática y aminoraba su velocidad de desplazamiento en el sentido del campo. Di-



BACTERIA QUE BUSCA EL SUR recogida por los autores y sus colegas cerca de Christchurch, Nueva Zelanda. Se observó que las bacterias magnetotácticas de Nueva Zelanda y de Tasmania nadaban predominantemente hacia el sur, de acuerdo con la hipótesis de los autores de que la magnetotaxia bacteriana era un mecanismo para emigrar hacia el fondo y permanecer en los sedimentos. En la mayoría de las especies recogidas en el hemisferio meridional, los magnetosomas tenían formas semejantes a las halladas en el septentrional. Pero fue en esta especie donde se descubrió una nueva configuración: puntiforme.

ríase, pues, que la velocidad de migración habría de depender de la fuerza del campo, en la forma que describiera Langevin, y, de cumplirse esto, el análisis de los datos proporcionaría el momento dipolar magnético de las bacterias. Experimentos recientes, realizados en Woods Hole por Kalmijn y sus colaboradores B. D. Teague y M. K. Gilson, sobre bacterias individuales, han confirmado la dependencia respecto del campo, en línea con la función de Langevin, y han calculado unos momentos magnéticos para cada bacteria coherentes con los datos de los análisis presentados más arriba.

El hecho de que las bacterias con movilidad unidireccional recogidas en sedimentos del hemisferio septentrional naden persistentemente hacia el norte se atribuye a la orientación, con respecto al aparato flagelar de la célula, del polo que busca el norte de su momento dipolar magnético. Si se invirtiera el momento magnético (es decir, si el polo que busca el sur se orientase con respecto al flagelo), las bacterias nadarían en dirección contraria (hacia el sur).

En el laboratorio de Kalmijn se han creado bacterias que buscan el sur. A sugerencia de Edward M. Purcell, de la Universidad de Harvard, se expusieron bacterias que buscaban el norte a breves e intensas ráfagas de un campo magnético que estaba orientado en sentido contrario al del campo ambiente. Cuando el nuevo campo inducido adquiría la intensidad suficiente, invertía el momento dipolar magnético de la cadena de magnetosomas. Las bacterias cuyo momento había sido invertido se orientaban ahora hacia el sur. Daban un giro en forma de U y nadaban en sentido contrario al inicial, es decir, en sentido contrario al del campo ambiente estacionario. Podía restablecerse su orientación en busca del norte mediante una segunda ráfaga antiparalela a la primera.

El momento magnético podía también redirigirse a través de un proceso tradicional de desimantación. Se exponía, primero, un grupo de bacterias buscadoras del norte a un campo magnético alterno de 60 hertz con una magnitud de pico de más de 1000 gauss, para ir luego retirándolas, lentamente, de dicho campo. Antes de la desimantación todas las bacterias nadaban hacia el polo norte; pero, consumada ésta, la mitad –más o menos– nadaba hacia el norte y la otra mitad se encaminaba hacia el sur. El experimento mostraba que, si bien podían invertirse los momentos dipolares magnéticos de las distintas bacterias, los organismos no se desimantaban, que era justamente lo mismo que cabía esperar de una cadena de partículas magnéticas de un solo dominio.

El comportamiento de los microorganismos, incluidos sus movimientos en respuesta a productos químicos o a la luz, constituye una ventaja adaptativa para su supervivencia. La magnetotaxia conferirá, presumiblemente, alguno de estos valores de supervivencia; pero ¿cuál? Para contestar la pregunta volvamos al modelo de Gilbert sobre la inclinación del campo magnético terrestre.

No está de más repetir que el campo geomagnético apunta al norte y hacia abajo en el hemisferio septentrional y al norte y hacia arriba en el hemisferio meridional, con un valor absoluto del ángulo de inclinación magnética que aumenta a medida que uno se desplaza desde el ecuador geomagnético hacia uno cualquiera de los polos. Debido a la inclinación del campo magnético terrestre, las bacterias que buscan el norte en el hemisferio septentrional migran hacia abajo y las bacterias que buscan el sur migran hacia arriba. En Nueva Inglaterra, la inclinación es de unos 70 grados hacia abajo respecto al norte; por tanto, la población de bacterias magnetotácticas que busca predominantemente el norte se dirigirá hacia abajo. Las bacterias se mantienen en los sedimentos y lejos de la superficie del agua.

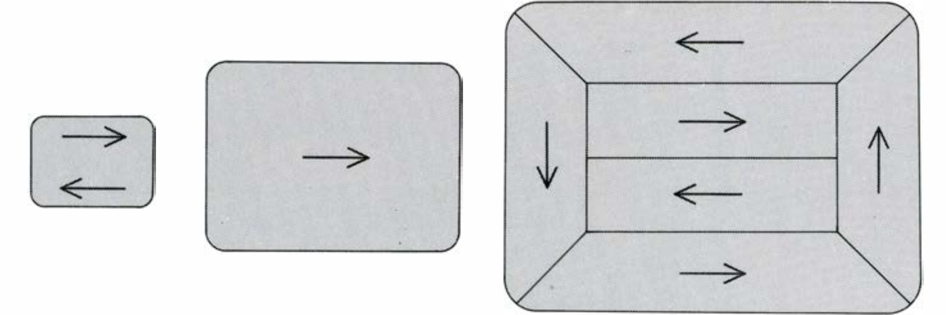
Las bacterias magnetotácticas son organismos del fondo, anaerobios (no toleran la presencia de oxígeno) o microaerófilos (que sobreviven mejor en un ambiente con poco oxígeno). Una tendencia a migrar hacia abajo les resultaría ventajosa en el sentido de que les ayudaría a evitar los efectos tóxicos de la mayor concentración de oxígeno de la superficie del agua.

Fiados, pues, en esta hipótesis, cabía esperar que, entre las bacterias magnetotácticas del hemisferio meridional, predominaran las buscadoras del sur que emigraran hacia abajo y permanecieran en los sedimentos y alejadas de la superficie del agua. En enero de 1981, junto con Nancy Blakemore y Kalmijn, emprendimos una expedición a Nueva Zelanda y Tasmania a la caza de bacterias magnetotácticas que nos validaran la hipótesis. Elegimos esos puntos en razón de su peculiaridad ecológica y por compartir muchas caracte-

ESPECTRO DE MÖSSBAUER del isótopo 57 del hierro hallado en células de *A. magnetotacticum* sometidas a criofractura (*curva coloreada*). El espectro coincide, casi puntualmente, con el de una muestra de magnetita pura (*curva negra*); ello demuestra que el hierro que hay en esta especie magnetotáctica se encuentra principalmente en la forma Fe_3O_4. Para medir el espectro se aplicaron rayos gamma con una energía de 14,4 keV (miles de electronvolt) a todas las muestras. Moviendo la fuente de rayos con respecto a la muestra, la longitud de onda de los rayos gamma podía desplazarse en proporción a su velocidad relativa. A determinadas velocidades relativas, los núcleos de hierro-57 de la muestra podían absorber los rayos gamma por transición a un estado excitado con una energía de aproximadamente 14,4 keV. El registro del espectro evidencia la intensidad relativa de los rayos gamma que golpean el detector en función de la velocidad de la fuente con respecto a la muestra. Las velocidades exactas a las que se observa cada absorción en una muestra revelan la forma química del hierro encerrado en esa misma muestra.

rísticas físicas con Nueva Inglaterra, tales como la latitud absoluta, la intensidad del campo magnético y el clima. Además, la inclinación magnética en Nueva Zelanda y en Tasmania tiene el mismo valor absoluto que en Nueva Inglaterra, aunque de signo contrario.

Se recogieron sedimentos de cienos salinos, fuentes termales, lagos de montaña y cavernas. Todos se examinaron con microscopios portátiles equipados con bobinas de Helmholtz. En la mayoría de los testigos había bacterias magnetotácticas y entre éstas, confirmando lo previsto, predominaban las buscadoras del sur. Al igual que sus homólogas del hemisferio septentrional, las bacterias tenían magnetosomas en su citoplasma y su polaridad podía invertirse mediante el procedimiento de desmagnetización. Concluyendo, pues, las bacterias magnetotácticas de ambos hemisferios son iguales en todo, salvo en la polaridad de las poblaciones.



EL TAMAÑO DE UNA PARTICULA de magnetita determina su momento magnético efectivo. La menor de las partículas, la representada a la izquierda, es superparamagnética: consta de un solo dominio magnético (una región donde los ejes de espín de todos los átomos de hierro están orientados en el mismo sentido), pero la orientación de su momento magnético (*flechas negras*) puede fluctuar como resultado de una disrupción térmica. La partícula mayor, a la derecha, tiene múltiples dominios magnéticos, donde los ejes de rotación están orientados de forma antiparalela; de acuerdo con esto, posee un bajo momento magnético residual, o neto. La partícula de tamaño intermedio, en el centro, mide entre 400 y 1200 angstrom de anchura. En este intervalo de tamaños, las partículas de magnetita son térmicamente estables, y tienen un solo dominio. Las partículas producidas por bacterias magnetotácticas se sitúan dentro de esta gama de tamaños intermedios, aprovechándose así al máximo las propiedades de la magnetita.

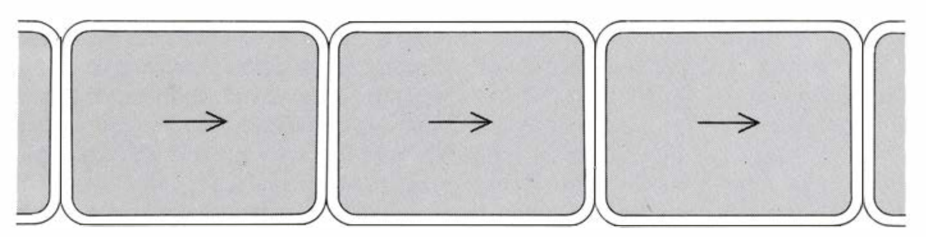
Los experimentos de laboratorio ponen de manifiesto también el papel que desempeña el signo de la inclinación magnética a la hora de seleccionar la polaridad dominante. En un sistema de bobinas eléctricas que invertía la componente vertical del campo magnético ambiental se colocaron muestras de sedimentos de Nueva Inglaterra, en donde las bacterias eran inicialmente buscadoras del norte. Las bobinas reproducían, pues, la orientación inclinada hacia arriba del campo geomagnético en el hemisferio Sur. Se introdujeron muestras control en bobinas que no llevaban corriente, y que por tanto permanecían en el campo magnético del ambiente, inclinado hacia abajo. Se examinaron las muestras periódicamente y se anotaron los datos relativos al número y polaridad de las células magnetotácticas, a lo largo de un año.

Mientras que las poblaciones control seguían dirigiéndose hacia el norte, la polaridad se invirtió en las poblaciones magnetotácticas expuestas al campo invertido. Seis días más tarde se detectaba ya el aumento del número de bacterias buscadoras del sur. Al cabo de tres semanas, constituían la fracción dominante, y, a las ocho, podíamos hablar ya de la inversión, casi absoluta, de la población. Resultado que concordaba de pleno con la hipótesis de que la componente vertical del campo geomagnético determinaba la polaridad de las poblaciones bacterianas en los entornos naturales. Del experimento se desprendía, además, otra conclusión: la respuesta a los cambios de la componente vertical transcurre a lo largo de muchas generaciones.

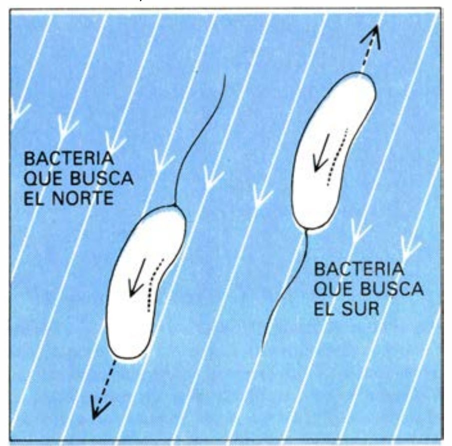
Una nueva confirmación de la importancia de la componente vertical del campo geomagnético procede de observaciones con bacterias magnetotácticas en el ecuador geomagnético, en donde la componente vertical es nula y las líneas del campo geomagnético son horizontales. Como todas las bacterias magnetotácticas están allí dirigidas horizontalmente, no hay polaridad hacia abajo. Por otro lado, el movimiento horizontal podría resultar ventajoso para las bacterias de los sedimentos, al reducir las excursiones hacia arriba, perjudiciales en potencia. Por supuesto, la magnetización de una u otra polaridad ofrecería idéntico interés y ninguna de las dos quedaría primada.

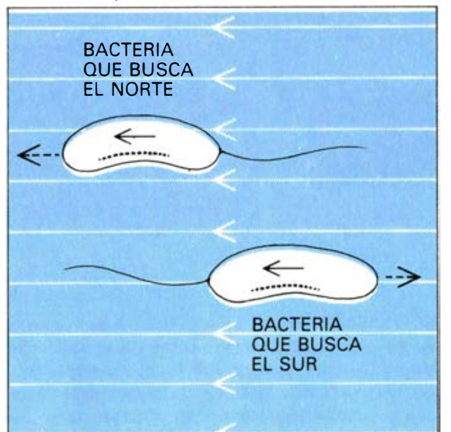
En los sedimentos procedentes de Fortaleza, en Brasil, lugar próximo al ecuador geomagnético, las bacterias magnetotácticas se dirigen tanto al norte como al sur. En esta observación se contó con la colaboración de F. F. Torres de Araujo, de la Universidad Federal de Ceará, y D. M. S. Esquivel y J. Danon, del Centro Brasileño de Investigación Física. En Río de Janeiro, donde la inclinación es de 25 grados hacia arriba, predominan las bacterias que buscan el sur. Queda por determinar la dependencia de las proporciones relativas de cada polaridad respecto de la inclinación magnética para inclinaciones entre cero y 25 grados.

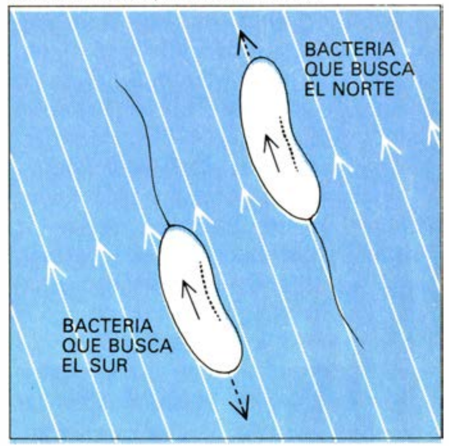
Wendy O'Brien, Randy Caplan y Nancy Blakemore han corroborado y ampliado, in vitro, estos hallazgos. Sedimentos procedentes de Nueva Inglaterra (en donde las bacterias buscan el norte) se colocaron en pequeños viales cerrados con tapones de goma algo permeables al oxígeno y se almacenaron en una caja con blindaje magnético. A lo largo de varias semanas, las muestras sufrieron cambios, resultando igual número de células buscadoras del norte que del sur. Las bacterias magnetotácticas de ambas polaridades podían observarse en la misma cantidad, en la superficie exhausta de oxígeno, del agua de los viales cerrados. Si se quitaban los tapones, permitiendo la entrada de más aire en los viales, no pasaban muchos días hasta que desaparecieran de la superficie del agua las bacterias magnetotácticas. Sí abundaban, por contra, bacterias de ambos tipos de polaridad entre el sedimento y el agua. Así pues, tanto los experimentos de campo como los realizados en el laboratorio muestran que, en ausencia de



LOS MOMENTOS MAGNETICOS están alineados en el mismo sentido en las cadenas de magnetosomas encontradas en la mayoría de las especies de bacterias magnetotácticas. El momento magnético total de una de estas cadenas es, por tanto, la suma de los momentos de las distintas partículas que la componen.







IMPORTANCIA DE LA MAGNETOTAXIA para la supervivencia de una bacteria que habita en los fondos. Se valora aquí en tres latitudes distintas. En el hemisferio septentrional (*izquierda*) una bacteria buscadora del norte se dirige hacia abajo, hacia los sedimentos, en tanto que una bacteria buscadora del sur nada hacia arriba, hacia la superficie del agua, más oxigenada (y por tanto, tóxica). En el meridional (*derecha*) las bacterias buscadoras del sur tienden a nadar hacia abajo y las buscadoras del norte hacia arriba. Cerca del ecuador geomagnético (*en el centro*) se dan ambos tipos, y uno u otro se desplazan horizontalmente, reduciéndose la potencialidad dañina de las excursiones hacia arriba. Las flechas blancas señalan el sentido del campo magnético.

una componente vertical del campo magnético, coexisten ambos tipos de polaridad. Los experimentos de laboratorio ponen de manifiesto, además, que algunos factores, pensemos en la concentración de oxígeno, influyen en la distribución de bacterias magnetotácticas en los ambientes naturales.

A. magnetotacticum ilustra la relación entre los distintos comportamientos de respuesta. La bacteria en cuestión posee un flagelo en cada uno de los dos extremos que le permiten nadar tanto hacia delante como hacia atrás. Así, una vez que las células se alinean pasivamente en el campo magnético terrestre, pueden nadar arriba y abajo a lo largo de las líneas del campo inclinado. La dirección elegida responde, al menos en parte, al gradiente local de la concentración de oxígeno.

Las células de este organismo microaerófilo son también aerotácticas: tienden a acumularse, en bandas, en las regiones donde hallan la concentración de oxígeno apropiada. La magnetotaxia refuerza la propiedad de aerotaxia, en el sentido de que reduce las excursiones al azar e induce una movilidad en línea recta.

Aunque la capacidad de sintetizar magnetita está codificada genéticamente, la polaridad de las bacterias magnetotácticas no puede estar determinada por el genoma. Cuando una bacteria sin magnetosomas empieza a fabricarlos, tiene igual probabilidad de ser buscadora del norte como de buscar el sur. Sin embargo, una bacteria que ya tiene magnetosomas puede transmitir la polaridad a sus células hijas, porque la cadena de magnetosomas se reparte durante la división de la célula. Si cada célula hija hereda algunos de los magnetosomas parentales, hereda la polaridad parental; a medida que la célula hija sintetiza nuevos magnetosomas en los extremos de su cadena heredada, éstos quedan magnetizados en la misma dirección en virtud de su interacción con las partículas imanadas de la cadena.

Ello no es óbice para que, en cada generación, aparezcan algunas bacterias de polaridad opuesta, quizá por no haber recibido ningún magnetosoma y haber sintetizado luego magnetosomas que fortuitamente han quedado magnetizados en sentido contrario. En circunstancias normales, las bacterias con magnetización invertida sufren una discriminación continua; se dirigen hacia arriba, a regiones menos adecuadas para su crecimiento, y la población de los sedimentos mantiene su polaridad original. Sin embargo, si la componente vertical del campo magnético se invierte, las que antes estaban mal dotadas se dirigen ahora hacia abajo y su descendencia acapara los sedimentos. De presentarse tal acontecimiento en un ambiente natural, la selección actúa en cuestión de semanas, si no antes.

Este experimento de laboratorio halla su réplica a escala global en las inversiones del campo magnético terrestre. No es este último estrictamente dipolar; ni se anulará tampoco, consiguientemente, en el curso de tal inversión. La realidad es que cambia de magnitud y rota a través de la inclinación cero en cualquier lugar de la superficie de la tierra durante un período de transición de varios millones de años. En el proceso, cabe presumirlo, se invertirá también la polaridad de las bacterias de ambos hemisferios.

La vida ha evolucionado en un campo magnético. Las bacterias magnetotácticas aportan la primera demostración inequívoca de que los organismos responden al campo geomagnético a través de una brújula biomagnética fabricada con el mineral inorgánico magnetita, que ellos sintetizan a partir del hierro soluble de su entorno. Brújula, que, en varios aspectos, parece hallarse bien adaptada para interaccionar con el campo magnético terrestre. Consta de partículas con un solo dominio magnético, portadoras cada una de un alto magnetismo residual y dispuestas frecuentemente en una cadena que se extiende a lo largo del eje de movilidad celular. Los magnetosomas no agotan su función en el cumplimiento de una misión natatoria; parece probable que desempeñen algún cometido más en la vida de las bacterias magnetotácticas. Tal vez sirvan de almacén del hierro sobrante o quizás ayuden a la célula a liberarse del peróxido de hidrógeno, un producto tóxico del metabolismo del oxígeno, que se descompone en presencia de hierro. Una y otra acción fisiológica son hoy objeto de estudio.

Pero el investigador tiene un campo todavía más prometedor: con los años ha ido apareciendo magnetita en muchos organismos, además de las bacterias. Por ahora la lista registra a los quitones (un molusco marino), abejas melíferas, mariposas, palomas mensajeras y, recientemente, delfines. La dilucidación del papel de la magnetotaxia bacteriana avala, pues, la búsqueda de un sentido magnético comparable en los organismos superiores. El resultado de esta investigación puede conferir una nueva realidad a la vieja idea del magnetismo animal.

Limitaciones de la fijación biológica de nitrógeno

Hay factores limitantes de la fijación biológica de nitrógeno que representan pérdidas importantes de energía para la planta, cuya superación puede contribuir a incrementar de forma considerable la cantidad de proteína disponible

Eulogio J. Bedmar y José Olivares

El aumento de la población mundial, el deterioro del medio ambiente y el agotamiento de los recursos fósiles no renovables han estimulado, entre otros factores, el interés por el estudio intensivo de los diferentes procesos implicados directamente en la producción agrícola, tales como la fotosíntesis y la fijación de nitrógeno. En un reciente estudio, la Academia Nacional de Ciencias de Estados Unidos ha previsto que, para proporcionar alimentos a la población del año 2000, se requerirán unos incrementos en el rendimiento agrícola de alrededor del 2,5 por ciento anual, a lo largo de los próximos 20 años, que permitan doblar la producción primaria de los mismos.

Aparte de otras posibilidades, más teóricas que prácticas, la disponibilidad de alimentos puede aumentarse por extensión del suelo dedicado a cultivo o por incremento del rendimiento por unidad de superficie. De acuerdo con la Organización Mundial para la Alimentación y la Agricultura (FAO), de los 3200 millones de hectáreas potencialmente cultivables, sólo se utilizan la mitad; el aprovechamiento de la otra mitad, aparte de exigir inmensos gastos para hacerla productiva, no sería económicamente rentable. No parece, por tanto, que la primera alternativa sea capaz de dar lugar a aumentos significativos, pues la tierra más adecuada está ya bajo cultivo, y, aún más, este suelo útil mengua continuamente en los países desarrollados por su empleo en otros usos, tales como viviendas, autopistas, industria, turismo y servicios. Estos hechos sugieren que el incremento por unidad de superficie es la única posibilidad que, por el momento, puede contribuir a solucionar el problema alimentario. Sin embargo, y a pesar de la existencia de una avanzada tecnología y de unos métodos adecuados de hibridación de plantas, que produjeron durante los años 60 y principios de los 70 resultados espectaculares en los rendimientos, el incremento medio en la mayoría de las cosechas, después de esta fecha, no ha aumentado de forma significativa [véase el número extraordinario y monográfico de *Investigación y Ciencia* dedicado a la alimentación y la agricultura, aparecido en noviembre de 1976].

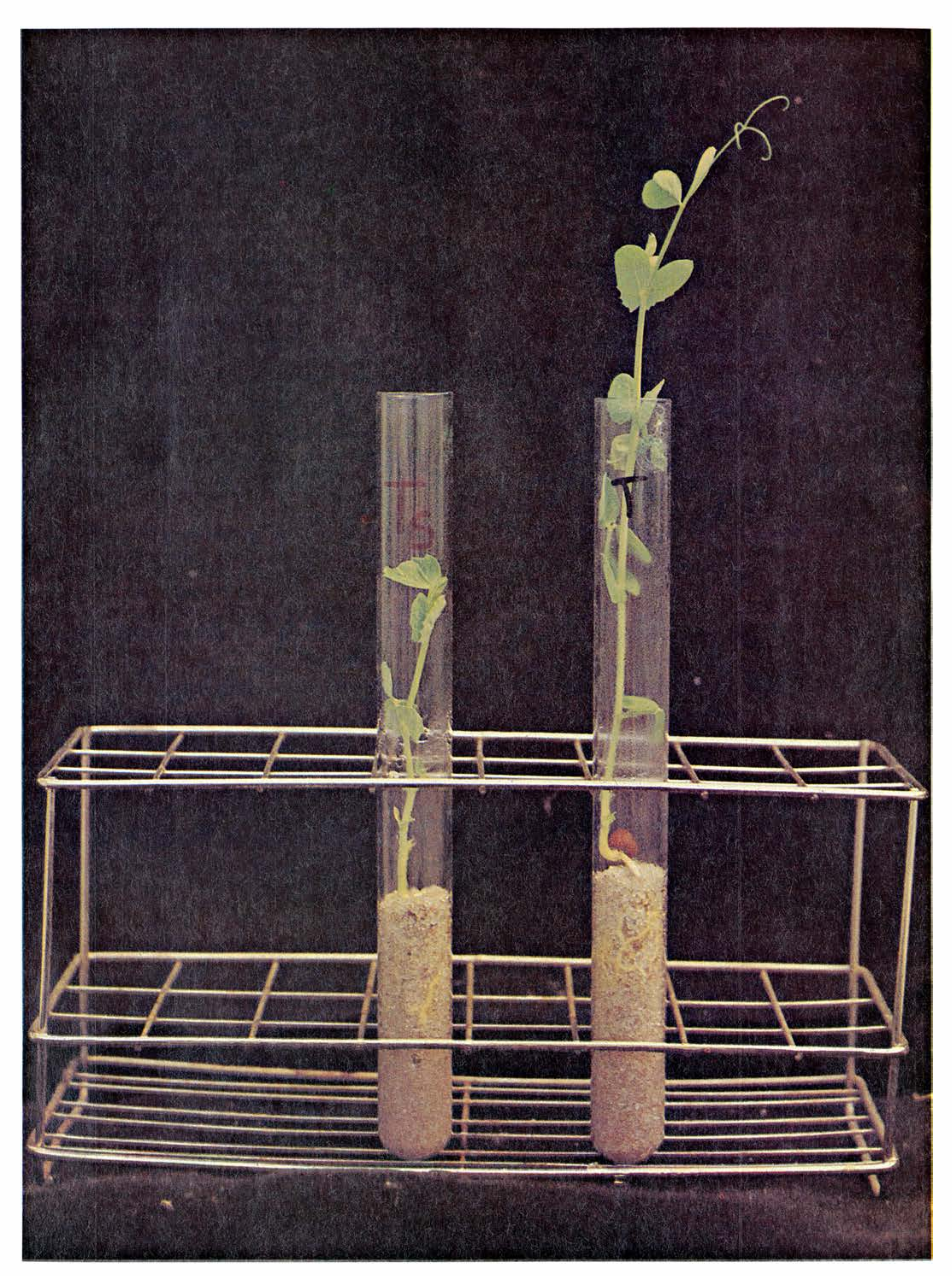
Entre los factores que han contribuido al incremento de la producción cabe destacar el empleo masivo de fertilizantes y, especialmente, los nitrogenados. Sin embargo, en la actual crisis energética, la progresiva carestía de estos productos, por el alto consumo de gas natural y petróleo utilizados en el proceso de su fabricación, obliga a buscar una alternativa que permita responder a la demanda de alimentos. En las circunstancias actuales, incluso los países más desarrollados se ven con dificultades para el empleo a gran escala de la fertilización nitrogenada.

A la relativamente baja eficiencia de la aplicación de nitrógeno, ya que la mitad del incorporado al suelo se pierde por desnitrificación y lavado, hay que añadir el peligro potencial de su uso, que está planteando graves problemas ecológicos derivados de la eutrofización de las aguas y del aumento de nitratos en el agua potable. Hay pruebas también de que la aparición de nitrosaminas en el ambiente, liberadas del ciclo del nitrógeno, encierra una amenaza potencial, y compuestos como el óxido de nitrógeno (N_2O), subproducto de la desnitrifación microbiana, pueden actuar como agentes destructivos de la capa de ozono de la atmósfera, protectora de las radiaciones.

Junto con el agua, el nitrógeno es el factor limitante más común para las plantas, a pesar de tratarse de uno de los elementos mayoritarios que constituyen los seres vivos y el componente mayoritario en la atmósfera. Este elemento, como tal, es inerte y solamente utilizable por las plantas y animales en forma fijada, mineral para las primeras y orgánica para los segundos. La sola presencia de nitrógeno en la atmósfera no asegura su asimilabilidad biológica; es necesario que, en forma combinada, pase al suelo para completar su ciclo dinámico supliendo las pérdidas que van sucediéndose por diferentes causas. La fijación biológica de nitrógeno es, por lo que supone de aporte de este elemento a la biosfera, de gran significación en la ecología de nuestro planeta y su potenciación aminoraría no sólo el gasto que representa la fertilización nitrogenada, sino también los mencionados peligros que su abuso lleva consigo.

Hay que tener en cuenta que del total anual de nitrógeno atmosférico fijado (210×10^6 toneladas según las últimas apreciaciones) aproximadamente 120×10^6 toneladas corresponden al fijado biológicamente. El resto se debe a la fijación puramente química o abiológica, bien de forma natural, por descargas eléctricas, combustión, ozonización y otros (30 millones de toneladas), o artificial, siguiendo el proceso Haber-Bosch (60 millones para 1977), que convierte el nitrógeno en amonio bajo condiciones de alta temperatura y presión, utilizando petróleo y gas natural como fuente de energía.

La fijación biológica es un proceso que, aunque muy extendido, aparece sólo en microorganismos procarióticos, libres o en asociación más o menos estricta con plantas. De todas las formas posibles, la de mayor significación agrícola es la fijación llevada a cabo por la simbiosis mutualística *Rhizobium*-leguminosa, ya que estas plantas son,



AUNQUE ES BIEN PATENTE el efecto de la inoculación sobre el crecimiento de una leguminosa en medio libre de nitrógeno, hay momentos en el desarrollo de la planta en los que la fijación de nitrógeno no puede soportar su correcta nutrición, como consecuencia de la falta de fotosintato disponible para el adecuado funcionamiento de la asociación simbiótica establecida (*Rhizobium*-leguminosa). La planta más crecida de la foto corresponde a la inoculada.

por su riqueza en nitrógeno, una fuente insustituible de proteínas para la alimentación humana y animal.

De acuerdo con las conclusiones del IV Congreso Internacional de Fotosíntesis (1977), la producción mundial de leguminosas grano, estimada en 130 millones de toneladas/año, debería incrementarse a lo largo de ésta y la próxima década hasta llegar, por lo menos, hasta los 500 millones de toneladas a final de siglo. Así se podría suministrar la proteína suficiente para las exigencias de la población del año 2000. Este reto requerirá mucho más esfuerzo que en el caso de los cereales, cuya producción se duplicó de 1950 a 1975 con sólo un 10 por ciento de aumento de la superficie cultivada. La producción de leguminosas, en cambio, sólo ha crecido el 15 por ciento correspondiente al procentaje de incremento del área en cultivo. Estas plantas, con una nutrición nitrogenada menos eficiente, con problemas de adaptación y sobre las que se ha desarrollado un escaso estudio genético, son, por lo general, menos rentables que los cereales. Por otra parte, la investigación realizada sobre éstos no ha tenido paralelismo en las leguminosas, lo que ha determinado el estancamiento de los rendimientos. Dado que estas plantas contribuyen en mayor proporción, entre todos los sistemas, a la incorporación de nitrógeno a la biosfera, se puede considerar constante la cantidad total de nitrógeno fijado en estos últimos 25 años.

En las circunstancias presentes, parece, pues, sensato prestar más atención a estas plantas, no sólo, como se ha indicado antes, por su particular incidencia en alimentación humana o animal, sino también por su relativa independencia de la fertilización nitrogenada. Tal especial comportamiento se debe a su capacidad para fijar nitrógeno molecular en simbiosis mutualística con *Rhizobium*. Esta asociación bacteria-planta utiliza, aunque indirectamente, la luz como fuente de energía para llevar a cabo la reducción del nitrógeno a amonio, lo que supone un ahorro importante de los combustibles fósiles necesarios para la fijación química por el proceso Haber-Bosch.

Sólo la luz puede permitir que la fijación biológica de nitrógeno por la simbiosis *Rhizobium*-leguminosa ocurra de forma eficaz, ya que se trata de un proceso altamente consumidor de energía. La reacción $N_2 + 3H_2O \rightarrow 2NH_3 + 3/2O_2$ requiere aproximadamente 340 kilojoule por mol de amonio fijado. Si la fijación tiene lugar acoplada con la oxidación de la glucosa se consumirán, en teoría, 1,44 gramos de azúcar por gramo de nitrógeno fijado. Pero la práctica enseña una proporción muy diferente. La raíz de una leguminosa inoculada con su apropiado *Rhizobium* utiliza aproximadamente 18,8 gramos de glucosa por gramo de nitrógeno fijado. Esta proporción se incrementa en los fijadores libres, llegando a ser en *Azotobacter* de 100 gramos de glucosa por gramo de nitrógeno. De aquí se puede deducir la baja significación agrícola de los fijadores libres, al no encontrar disponible en el medio tan alta cantidad de materia orgánica utilizable.

En leguminosas, aproximadamente cuatro moléculas de adenosintrifosfato (ATP) son hidrolizadas por cada dos electrones transferidos por la nitrogenasa, enzima responsable de la reducción de nitrógeno (N_2) a amonio (NH_3), primer producto estable que resulta de la fijación. A estas 12 moléculas de ATP hay que sumarles otras nueve, consumidas en la provisión de los tres pares de electrones necesarios para la reducción. De esta forma, sólo en el paso de N_2 a $2NH_3$ se requiere un gasto teórico de 21 moléculas de ATP. Si se considera un rendimiento del 80 por ciento se puede suponer que la fijación de una molécula de N_2 consume unas 24 moléculas de ATP, a las que hay que añadir las requeridas para la síntesis de la nitrogenasa, que manifiesta un elevado "turnover", como corresponde a un enzima sometido a una estricta regulación.

La nitrogenasa acopla la hidrólisis de ATP con la transferencia de electrones procedentes de un donador reducido, ferredoxina o flavodoxina, para pasar el N_2 a NH_3. Los factores primarios que podrían limitar esta reacción son: concentración de nitrogenasa, grado de saturación de la misma por los substratos de la reacción y la concentración del producto final, amonio. De las pruebas se deduce que la concentración de nitrogenasa no es, por ella misma, factor limitante; el bajo valor de su constante de equilibrio para el nitrógeno (de 0,02 a 0,05 atmósferas) elimina la presión parcial de éste como una limitación práctica en condiciones normales. (La constante de equilibrio, o Km, es un valor de la cinética de las reacciones enzimáticas.)

El amonio es un represor de la síntesis de la nitrogenasa a nivel de transcripción del ADN correspondiente, de forma que en presencia de cierta concentración de amonio no se sintetiza el enzima. En *Rhizobium*, tal tipo de inhibición, universal para todos los fijadores e incluso para esta misma bacteria en condiciones de vida libre, no ocurre cuando se encuentra en simbiosis con la leguminosa correspondiente. El amonio formado se exporta casi en su totalidad, hasta la célula vegetal hospedadora, donde lo metabolizan enzimas adecuados (glutamina sintetasa y glutamato sintetasa) y lo transporta el xilema al resto de la planta.

En las leguminosas, *Rhizobium* utiliza directamente el fotosintato (productos carbonados formados durante la fotosíntesis) para llevar a cabo, en simbiosis, el proceso fijador. Esta función depende, por tanto, del suministro de substrato aprovechable que no sólo es fuente de energía, sino que, además de soportar el crecimiento del nódulo (manifestación visible de la infección de la leguminosa por *Rhizobium* y lugar donde se lleva a cabo la fijación), proporciona el poder reductor necesario y provee del esqueleto carbonado adecuado para la incorporación del amonio formado.

FACTORES	INCREMENTAN LA FIJACION	DISMINUYEN LA FIJACION
BACTERIA	RAZA	
PLANTA	VARIEDAD	
FOTOSINTESIS	ALTA	BAJA
FOTORRESPIRACION	BAJA	ALTA
HIDROGENASA	ALTA	BAJA
FERTILIZANTE	FOSFATOS	NITRATOS-AMONIACO
DENSIDAD DE SIEMBRA	BAJA	ALTA

FIJACION DE NITROGENO por la simbiosis *Rhizobium*-leguminosa. Se ve afectada positiva o negativamente por numerosos factores, dependientes de la bacteria, de la planta o de la interacción bacteria-planta. Aunque cada leguminosa o grupo de ellas forma simbiosis mutualística con una determinada especie de *Rhizobium*, hay, para una especie bacteriana determinada y su correspondiente leguminosa, razas y variedades, respectivamente, que constituyen asociaciones más eficaces que otras. Se trata de las que hay que seleccionar, en razón de su utilidad práctica, en la preparación de inoculantes microbianos.

Todavía no es posible especificar cuál de los papeles que representa el fotosintato es el más crucial. Los requerimientos de ATP para el funcionamiento de la nitrogenasa, así como para las reacciones de incorporación del amonio producido, podrían sugerir que la principal función del fotosintato es la producción de energía (en forma de ATP) y de poder reductor. Sin embargo, todo parece indicar que no hay prioridades, dada la interrelación existente entre las distintas posibles utilizaciones del fotosintato disponible.

De las numerosas limitaciones que tiene la fijación de nitrógeno, unas derivadas de la planta (resistencia a la infección e incapacidad para el establecimiento de la simbiosis), otras propias del microorganismo (problemas de reconocimiento de la planta hospedadora, baja infectividad y baja efectividad) y unas terceras del conjunto planta-bacteria, hay dos de especial importancia por su incidencia en el rendimiento del proceso: adecuado suministro de fotosintato y la presencia de actividad hidrogenasa capaz de utilizar el hidrógeno liberado en la fijación.

SISTEMA	ACTIVIDAD NITROGENASA PERDIDA COMO LIBERACION DE H_2 PORCENTAJE
PLANTAS LEGUMINOSAS	
ALFALFA	48
TREBOL	50
ALTRAMUZ	42
SOJA	55
COWPEA	<1
NO LEGUMINOSAS	
ALISO (*ALNUS*)	6
MYRICA SP.	2
ELEAGNUS SP.	13

LA PRESENCIA DE ACTIVIDAD HIDROGENASA en los fijadores libres o simbióticos es variable. Dado que en la fijación de nitrógeno siempre hay liberación de hidrógeno, cuando no existe una actividad hidrogenasa alta, parte de la energía disponible se pierde y, por tanto, parte de la capacidad fijadora.

A Ralph Hardy y colaboradores, de Du Pont de Nemours and Company (Wilmington, Delaware), se debe una demostración evidente de que la fotosíntesis constituye el principal factor limitante de la fijación: los resultados obtenidos del cultivo de soja en ambientes enriquecidos en CO_2 (1000 partes por millón en lugar de las 300 ppm de la atmósfera normal). La cantidad de nitrógeno fijado en estas condiciones aumentaba de 75 a 425 kilogramos por hectárea, lo que quiere decir que las plantas fijaron el 85 por ciento de sus requerimientos en nitrógeno, mientras que las no tratadas sólo consiguieron el 25 por ciento. El CO_2 adicional fue proporcionado continuamente desde el período de floración hasta el de senectud. No es práctico aumentar el nivel de CO_2 alrededor de las plantas, ya que se difundiría rápidamente al aire y no sería aprovechable de modo directo por las mismas, salvo en el caso de ambientes cerrados. Si tenemos en cuenta que el rendimiento medio del proceso fotosintético no es muy alto, debido a que sólo es aprovechable un bajo porcentaje del total de la energía luminosa que llega a la planta, cualquier alteración de esta eficiencia, en uno u otro sentido, repercutirá directamente en los niveles de fijación de nitrógeno alcanzados.

La fotorrespiración, aparente inversión del metabolismo carbonado fotosintético, resulta de la absorción de O_2 en presencia de la luz y liberación de CO_2. En determinadas circunstancias, esta inversión reduce considerablemente el potencial energético obtenido por la fotosíntesis, mengua que puede llegar hasta el 50 por ciento en algunas plantas de marcada importancia económica. Esto supone una seria reducción de la fotosíntesis neta y, por tanto, de la cantidad de fotosintato disponible para la fijación en el caso de las leguminosas.

Recientemente, se ha ido acumulando pruebas sobre pruebas que apoyan la sospecha de que el primer paso del proceso fotorrespiratorio lo constituye la síntesis de fosfoglicolato, vía ribulosa-1,5-bisfosfato, catalizada por el mismo enzima que lleva a cabo la fijación inicial de CO_2, la ribulosa-1,5-bifosfato carboxilasa, presente en el estroma de los cloroplastos. En 1971, William L. Ogren y sus colegas del Laboratorio USDA de la Universidad de Illinois, encontraron que ese enzima catalizaba la reacción de incorporación de O_2, por lo que se le conoce también como ribulosa-1,5-bifosfato oxigenasa-carboxilasa. Posee un peso molecular de 560.000, está constituida por ocho subunidades grandes, de peso molecular 58.000, y otras ocho menores de p.m. 15.000 y lo activan pequeñas concentraciones de CO_2 junto con iones magnesio.

Las subunidades de elevado peso molecular, codificadas por el genoma del cloroplasto, se sintetizan en su interior; las de menor peso molecular, por contra, son codificadas por el núcleo y sintetizadas por los ribosomas citoplasmáticos. El precursor de estas subunidades pequeñas cruza la envoltura del cloroplasto y se combina en el estroma con las subunidades de mayor tamaño para constituir el enzima completo.

El CO_2 compite con el O_2 como sustrato en condiciones favorables para éste, debido a las diferencias de concentración de ambos en la atmósfera normal: 20 por ciento de O_2 frente a 0,03 por ciento de CO_2. Por cada mol de CO_2 fijado en la reacción carboxilásica se producen dos moles de fosfoglicerato, mientras que en la reacción oxigenásica se originan, por cada mol de O_2 asimilado, un mol de fosfoglicerato y un mol de fosfoglicolato, que, una vez desfosforilado por una fosfatasa específica, abandona el cloroplasto. En los peroxisomas, el glicolato es oxidado por la glicolato oxidasa, enzima marcador de la actividad fotorrespiratoria, a glioxalato que, posteriormente, es aminado a glicocola. Y en la mitocondria, dos moléculas de glicocola forman, con desprendimiento de dióxido de carbono, una molécula del aminoácido serina que, vía hidroxipiruvato y glicerato, retorna al cloroplasto para ser fosforilada a fosfoglicerato e incorporarse de nuevo al ciclo.

Este tránsito consume energía y, por este mecanismo, la fotorrespiración puede llegar a producir pérdidas de hasta el 50 por ciento del CO_2 asimilado, mostrándose así como un mal inevitable, consecuencia perjudicial de una fotosíntesis que ha de realizarse en una atmósfera con limitada concentración de CO_2. Cuando se hace descender artificialmente a niveles aún más bajos, se puede llegar a una situación, conocida por punto de compensación de CO_2, en la que el carbónico eliminado por fotorrespiración está equilibrado con el asimilado por fotosíntesis. En

tal caso no hay ni pérdida ni ganancia de materia orgánica.

Cabe preguntarse si esta pérdida de energía a consecuencia de la fotorrespiración tiene alguna significación fisiológica y comporta, en contrapartida, algún beneficio para la planta. Independientemente de la hipótesis que la considera como una inútil reliquia evolutiva que la planta tiene que soportar, parece más acertada la idea de que se trata de una forma de disipación de la energía atrapada por el aparato fotosintético, en unas condiciones en que no cabe fotosíntesis normal. Es lo que ocurriría, por ejemplo, en una situación de stress de agua en que los estomas se cierran e impiden la difusión del CO_2. Por este mecanismo se protege así a la planta de un posible daño por fotooxidación. En tales circunstancias, el CO_2 asimilado fotonsintéticamente es desprendido por fotorrespiración, para volver a asimilarse, y así sucesivamente, sin que deje de funcionar la maquinaria fotosintética.

Como se ha indicado arriba, la síntesis y subsiguiente oxidación del glicolato es la causa de la fotorrespiración y, por tanto, de que muchas plantas ofrezcan escaso rendimiento fotosintético. Entre éstas se encuentran las leguminosas, que pertenecen a las plantas llamadas tipo C_3, por ser una molécula con tres átomos de carbono la primera en producirse en el proceso fotosintético.

SOLO UN BAJO PORCENTAJE de la energía luminosa que llega a las hojas es aprovechada por la planta. A pesar de ello, parte es malgastada en la mayoría de los casos (plantas C_3), en la fotorrespiración y, en leguminosas, en la liberación de hidrógeno por los nódulos. La distribución de la energía absorbida en una leguminosa se ha esquematizado en esta figura donde se puede apreciar que parte de ella retorna del nódulo al resto de la planta en forma de compuestos carbonados. Las flechas en color más intenso corresponden a las pérdidas correspondientes por los conceptos mencionados. (Ilustraciones de M. Alonso.)

Hasta mediados de los años 60 se pensaba que todas las plantas poseían igual mecanismo fotosintético. En 1965, Hugo Kortschak y colaboradores, de la Hawaiian Sugar Planters Association, y M. D. Hatch y colegas, de la Colonial, Sugar Refining Company, en Queensland, Australia, comprobaron que, en caña de azúcar, sorgo, mijo, maíz y otros, el producto inicial de la fotosíntesis era una molécula con cuatro átomos de carbono, el oxalacetato (AOA) formado por incorporación de CO_2 al fosfoenolpiruvato con intervención del enzima fosfoenolpiruvato carboxilasa, presente en el mesófilo de la hoja. El AOA es transformado en aspartato o malato que la planta difunde a través de los plasmodesmos hasta las células adosadas al parénquima vascular, donde ocurre la descarboxilación de estos ácidos; el CO_2 liberado se incorpora entonces con la intervención de ribulosa-1,5-bifosfato carboxilasa oxigenasa, como expusimos antes, para seguir el ciclo de Calvin. A la especial disposición de los tejidos foliares de estas plantas se le denomina estructura Kranz y, C_4, al tipo de fotosíntesis que realizan.

Se puede afirmar, que en las plantas con fotosíntesis C_4 se da una separación espacial de ambas reacciones de carboxilación. Hay que tener en cuenta que la fosfoenolpiruvato carboxilasa (presente en las células externas del mesófilo) muestra una afinidad por el CO_2 mayor que la ribulosa-1,5-bifosfato carboxilasa (confinada en las células internas del mesófilo). Esta adaptación permite mantener una elevada concentración de CO_2 en el lugar donde ocurre la reacción catalizada por la ribulosa-1,5-bifosfato carboxilasa, si bien este sistema de bombeo del CO_2 exige un consumo de ATP (cinco moléculas por cada una de CO_2 fijado) mayor que las plantas tipo C_3, que sólo requieren tres moléculas de ATP.

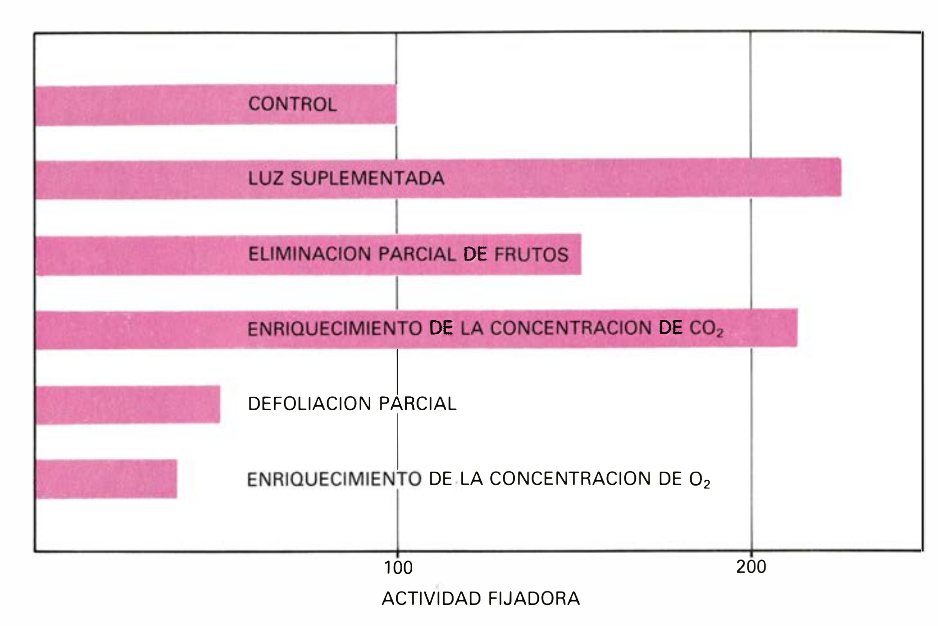
Una ventaja de las plantas con fotosíntesis C_3 estriba en la habilidad de su mecanismo para rendir mucho más a temperaturas de frías a moderadas y en condiciones de baja luminosidad, lo que las hace más idóneas para las zonas templadas. James R. Ehleringer y D. Björman, de la Universidad de Utah en Salt Lake, demostraron, en 1977, que la relación entre la actividad oxigenasa y la actividad carboxilasa aumentaba cuando la temperatura se elevaba de 15 grados Celsius a 35 grados C. Hay que esperar, por tanto, una disminución de la productividad en las plantas C_3 al au-

mentar la temperatura por incrementarse la fotorrespiración.

Entre las ventajas que presentan las plantas con fotosíntesis de tipo C_4 cabe señalar una mayor eficacia en el uso del agua, una menor tasa de transpiración y, sobre todo, un mayor aprovechamiento del nitrógeno: las plantas C_4 producen más fotosintato por gramo de nitrógeno que las plantas C_3. Estos hechos hacen que aquéllas puedan desarrollarse mejor en hábitats cálidos y áridos. Por otra parte, estas plantas presentan tasas de fotorrespiración apenas detectables, por lo que son más productivas que las C_3 bajo las condiciones descritas.

Precisamente en la rizosfera de plantas de este tipo Döbereiner, descubrió en Brasil niveles de fijación de nitrógeno relativamente importantes. Dos especies vegetales, *Paspalum notatum* y *Digitaria decumbens*, formaban con *Azotobacter paspali* y *Spirillum lipoferum* (*Azospirillum brasiliense*), respectivamente, lo que hoy se conoce como simbiosis asociativa. Estos microorganismos fijan nitrógeno utilizando como substrato las sustancias nutritivas exudadas por la raíz de las plantas correspondientes. Si bien del sistema radical de todas las plantas se liberan aminoácidos, vitaminas, azúcares y otros metabolitos en mayor o menor cantidad, el exudado radical de las plantas C_4 es particularmente rico y abundante, lo que permite la implantación y desarrollo de microorganismos que le son beneficiosos. Tales bacterias no forman estructuras morfológicas especiales. A lo sumo, penetran ligeramente en la lámina media de las células corticales y pueden fijar nitrógeno cuando crecen en ausencia de la planta. Por ello, la asociación que forman no es una auténtica simbiosis mutualística, sino una simbiosis asociativa o rhizocoenosis diazotrófica, tema de profundo estudio en la actualidad con vistas a un aprovechamiento más general e intensivo.

Cualquier intento dirigido a eliminar o disminuir el nivel de fotorrespiración en las plantas C_3 redundará en un aumento de la fotosíntesis neta que, en el caso de las leguminosas, se traducirá en una potenciación de la fijación de nitrógeno, tan estrechamente ligada a aquélla. En 1974, Israel D. Zelitch, de la Estación Experimental Agrícola de New Haven, encontró, en experimentos realizados con hojas de tabaco bajo condiciones de laboratorio, que el glicidato (2,3-epoxipropionato), un análogo del glicolato y del glioxalato, inhibía, en un 50 por ciento, tanto la síntesis del glicolato como la fotorrespiración. Posteriormente, pudo probarse que el glicidato inhibía la conversión del glioxalato en glicocola. En 1977, David J. Oliver, del mismo laboratorio, demostró que cuando se mantenían discos de hojas de tabaco en soluciones de aspartato, glutamato o glioxalato, bajo iluminación artificial, se inhibía el proceso fotorrespiratorio. No se conoce todavía el mecanismo por el cual, al aumentar la concentración de estos metabolitos en los tejidos de la hoja, queda bloqueada la ruta del glicolato.



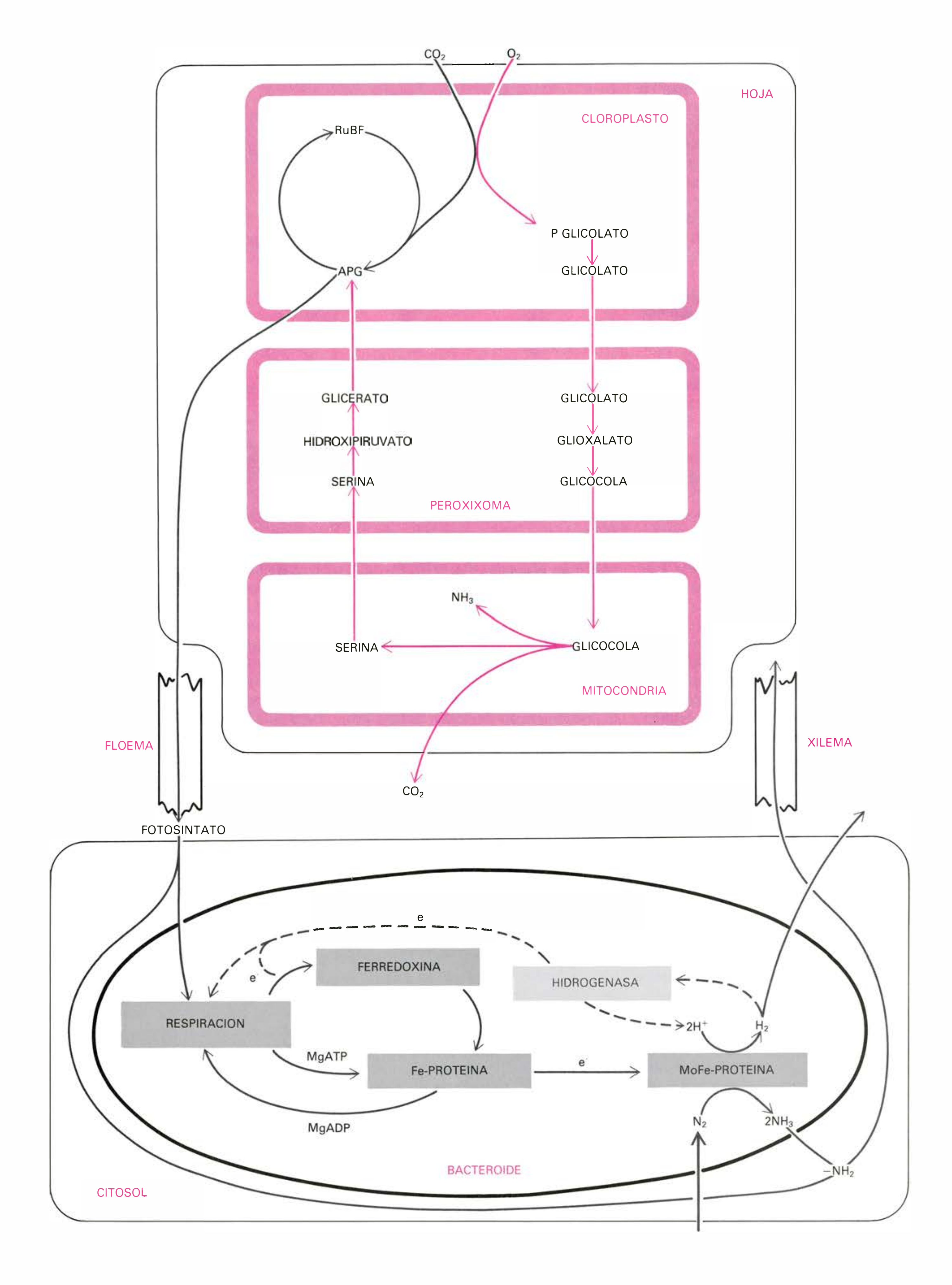
LA ACTIVIDAD FIJADORA, medida por la técnica de reducción de acetileno a etileno, es muy distinta cuando las plantas se someten a diferentes condiciones que afectan directa o indirectamente la cantidad de fotosintato (productos carbonados que se han formado durante el proceso de la fotosíntesis) disponible.

Se han empleado otros compuestos químicos para disminuir la fotorrespiración por obstrucción de algún paso del metabolismo del glicolato, si bien estos ensayos se han llevado siempre a cabo con trozos de hojas de diversas plantas o en cloroplastos aislados. R. Chollet, actualmente en el Laboratorio de Bioquímica Agrícola de la Universidad de Nebraska, fue el primero en comprobar que, tratando cloroplastos aislados de espinaca con glicidato, aumentaba la tasa de fotosíntesis neta hasta un 60 por ciento en virtud de la inhibición de la fotorrespiración.

En el Departamento de Microbiología de la Estación Experimental del Zaidín, en Granada, se están acometiendo, por los autores de esta revisión, una serie de experimentos dirigidos a comprobar el efecto de los inhibidores de la fotorrespiración sobre la fijación de nitrógeno en plantas de alfalfa noduladas. La aplicación, en el momento adecuado del desarrollo vegetativo de estas plantas, de soluciones de distinta molaridad de glioxalato, hidracida del ácido isonicotínico y glicidato, produce un incremento significativo de la actividad nitrogenasa de los nódulos de las raíces de las plantas tratadas respecto a los controles.

Los incrementos en fotosíntesis neta obtenidos con los tratamientos de tipo químico podrían igualmente conseguirse por procedimientos de tipo genético. Peter A. Day e Israel D. Zelitch observaron en 1968 que una variedad de tabaco, deficiente en clorofila y de lento crecimiento, presentaba tasas de fotosíntesis neta menores y de fotorrespiración más elevada que la variedad silvestre. Estos mismos autores encontraron, accidentalmente, otra variedad de tabaco, la Havana Seed, que poseía la mitad de la tasa normal de fotorrespiración. Por autopolinización de estas plantas se obtuvo una descendencia en la cual el 25 por ciento de los individuos presentaban niveles de fotosíntesis neta un 50 por ciento más alta de lo normal. Proporción que no aumentaba, sin embargo, en las sucesivas generaciones. Estos hechos demostraban que ciertos cambios genéticos podían controlar la fotorrespiración, así como el crecimiento en determinadas plantas.

Mary A. Berlyn e Israel D. Zelitch han realizado un amplio muestreo de poblaciones celulares utilizando cultivos de tejidos irradiados con luz ultravioleta para obtener líneas celulares resistentes a concentraciones de hidracida del ácido isonicotínico que inhiben la conversión de glicocola en serina.



CO_2
O_2
HOJA
CLOROPLASTO
RuBF
APG
P GLICOLATO
GLICOLATO
GLICERATO
HIDROXIPIRUVATO
SERINA
GLICOLATO
GLIOXALATO
GLICOCOLA
PEROXIXOMA
NH_3
SERINA
GLICOCOLA
MITOCONDRIA
FLOEMA
XILEMA
CO_2
FOTOSINTATO
e
FERREDOXINA
HIDROGENASA
e
RESPIRACION
$2H^+$
H_2
MgATP
Fe-PROTEINA
e
MoFe-PROTEINA
MgADP
N_2
$2NH_3$
BACTEROIDE
$-NH_2$
CITOSOL

Las células de plantas procedentes de estas líneas celulares mantienen su resistencia, lo que indica la estabilidad y la transmisión genética de tal carácter.

Hay, pues, claros indicios de que puede incrementarse la fotosíntesis neta en las plantas tipo C_3 inhibiendo la fotorrespiración por medios químicos o genéticos. Esta modificación daría lugar a un mayor rendimiento de las cosechas de marcado interés agrícola y económico, especialmente de las leguminosas, donde tendría lugar una fijación de nitrógeno más eficiente.

La limitación de la fijación impuesta por la disponibilidad de fotosintato queda patente cuando la planta necesita un suministro mayor de productos carbonados, como ocurre en la fase de formación de la semilla. El drenaje de fotosintato hacia el fruto es tal que la actividad fijadora en los nódulos decae drásticamente, comenzando su degeneración. No hay que ponderar la importancia que en esos momentos tiene una fotorrespiración baja o inhibida. Hasta ahora, en la práctica, esta disminución de la fijación sólo puede paliarse por la aplicación al cultivo de fertilizantes nitrogenados de rápida absorción. Se consigue así, incluso a los precios actuales del nitrógeno fijado químicamente, rendimientos económicos de más del 25 por ciento. Sin embargo, la fertilización ha de realizarse en la época justa para evitar el bloqueo de la fijación biológica si todavía se está llevando a cabo. No hay que olvidar que el nitrógeno combinado, aplicado a las leguminosas indiscriminadamente en tiempo y cantidad, perjudica, más que beneficia, la marcha del cultivo y puede ser causa de graves daños económicos.

La nitrogenasa, enzima fundamental de fijación de nitrógeno, es relativamente inespecífica. Además del N_2 a NH_3, puede reducir numerosos compuestos con triples enlaces. En ausencia de otros aceptores, el H^+ se reduce también a H_2; la verdad es que siempre se libera algo de H_2 aún en presencia de nitrógeno. La limitación que supone una actividad hidrogenasa baja fue puesta de manifiesto, en 1976, por Karel R. Schubert y Harold J. Evans, del Departamento de Botánica y Fitopatología de la Universidad del estado de Oregón, al comprobar que, durante el proceso de fijación, se perdía ATP como consecuencia de la evolución de hidrógeno catalizado por una actividad hidrogenasa, concomitante a la actividad nitrogenasa. La existencia de una actividad hidrogenasa en los nódulos de las raíces de las leguminosas que sea capaz de reutilizar el H_2 liberado por la nitrogenasa es el segundo de los factores importantes que determina la eficacia de la fijación de nitrógeno.

Por cada molécula de H_2 producido se hidrolizan unas cuatro moléculas de ATP. Ahí reside un importante problema, ya que esta reacción colateral de la nitrogenasa es causa de bajo rendimiento en los sistemas fijadores, incluidos los simbióticos, pues se malgasta un alto porcentaje de energía (entre el 40 y el 60 por ciento) en una función sin interés aparente.

En el campo de las relaciones entre hidrogenasa y nitrogenasa R. O. D. Dixon, del Departamento de Botánica de la Universidad de Edimburgo, realizó el esfuerzo más importante hasta la fecha, utilizando nódulos de guisantes. Dixon comprobó la existencia de dos sistemas de hidrogenasa independientes, uno de los cuales estaba implicado en la toma de H_2 y otro en su liberación. Esto es, funcionaban en distinto sentido, pero mientras que uno era sensible al CO, el otro no. Por otra parte, uno catalizaba la reacción $2H^+ \rightarrow H_2$ (nitrogenasa), mientras que el otro era reversible en la mayoría de los microorganismos, aunque se han descrito hidrogenasas en algunos fijadores que sólo actúan en la dirección $H_2 \rightarrow 2H^+$.

De acuerdo con Dixon, se puede considerar que la hidrogenasa desempeña una serie de funciones en los nódulos: mantener la nitrogenasa en el ambiente anaerobio adecuado (por eliminación del O_2 que se emplea en oxidar el hidrógeno liberado), impedir la inhibición de la nitrogenasa por el hidrógeno y, sobre todo, disminuir el balance final de ATP requerido para la liberación del H_2, ya que al reciclarse se producen nuevas moléculas de ATP recuperándose, en parte, las malgastadas por dicha liberación.

Durante la catálisis por la nitrogenasa, los protones y el nitrógeno compiten por los electrones de manera que, en ausencia de nitrógeno, el flujo de energía utilizable para la actividad nitrogenasa se emplea en producir hidrógeno. En presencia de nitrógeno, la evolución de H_2 ocurre simultáneamente a la formación de NH_3 y, durante el proceso, del 25-30 por ciento del total de equivalentes de poder reductor aprovechable por la nitrogenasa es dirigido hacia la reducción de H^+. En algunos casos, la pérdida de energía por esta vía puede considerarse equivalente a la debida a la fotorrespiración. Los bacteroides presentes en los nódulos que manifiestan una hidrogenasa activa pueden aprovechar este hidrógeno, bien oxidándolo a través de la cadena transportadora de electrones o utilizándolo como fuente de poder reductor para la nitrogenasa.

Dixon comprobó que al inocular diferentes leguminosas con una misma raza de *R. leguminosarum* se producían nódulos con diferentes niveles de actividad hidrogenasa, lo que indica que la leguminosa desempeña un importante papel en la expresión de este enzima.

Puesto que la evolución del hidrógeno es un proceso que depende del ATP, y que puede competir con la fijación de nitrógeno por la energía disponible, Schubert y Evans examinaron diferentes leguminosas para conocer la existencia y magnitud de la liberación de hidrógeno por la asociación bacteria-planta establecida. Así, llegaron a definir la eficiencia relativa (ER) de los nódulos fijadores de nitrógeno, considerada como la fracción del flujo total de electrones a través del sistema nitrogenasa que es aparentemente utilizado en el proceso de reducción del nitrógeno. De esta manera:

$$ER = 1 - \frac{H_2 \text{ liberado (en aire)}}{\text{Acetileno reducido}}$$

Para el cálculo de la ER hay que tener en cuenta que, bajo atmósfera de aire, la energía y el poder reductor suministrado por el fotosintato se reparte entre la reducción de nitrógeno y la liberación de hidrógeno. En cambio, en ausencia de nitrógeno (cuando éste es sustituido por argon), el consumo total de energía es proporcional a la tasa de reducción de protones, que se puede medir por técnicas amperométricas.

A niveles saturantes de acetileno, el total de energía usado por la nitrogenasa es proporcional a la tasa de reduc-

FOTOSINTESIS Y FIJACION DE NITROGENO son procesos biológicos que guardan una estrecha relación. El fotosintato que llega al nódulo a través del floema es utilizado como fuente de energía, de poder reductor y de esqueleto carbonado para la incorporación del amonio formado. Esta incorporación ocurre en el citosol de la célula hospedadora y los aminoácidos sintetizados son transportados por el xilema al resto de la planta. La fotorrespiración (*circuito de color rojo*) puede disminuir de forma considerable el nivel de fotosintato disponible. No obstante, la energía que se pierde en la formación de hidrógeno se puede recuperar, en cierto grado, si existe actividad enzimática hidrogenasa en los bacteroides.

ción del acetileno a etileno (medida por cromatografía gaseosa). El acetileno es uno de los sustratos de la nitrogenasa y la técnica de medida de la reducción de acetileno a etileno se utiliza extensamente para la determinación de la actividad nitrogenasa en raíces noduladas, nódulos aislados o microorganismos fijadores libres, por su simplicidad, precisión y rapidez. Cuando hay acetileno presente a concentración suficiente, toda la energía y poder reductor se utiliza en su reducción, no habiendo ni fijación de N_2 ni liberación de hidrógeno. Por último, hay que tener en cuenta que el flujo de energía aprovechable por la nitrogenasa no se altera por cambios en el ambiente gaseoso.

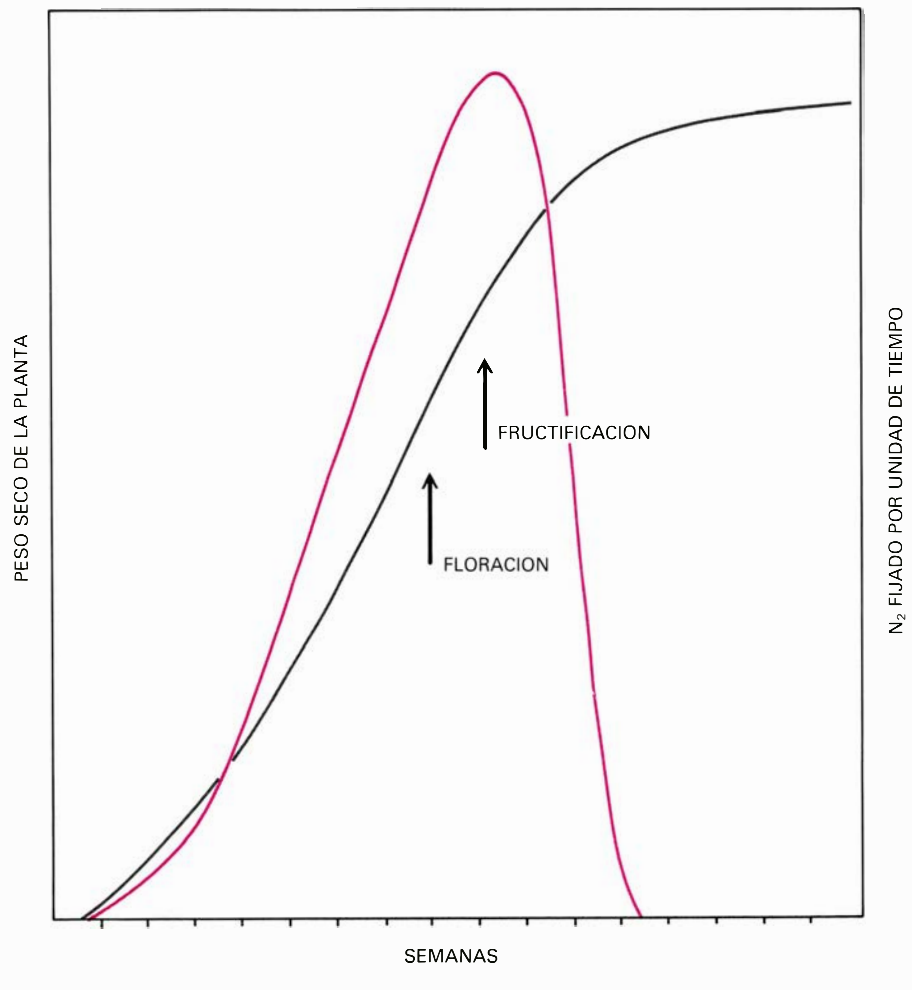
La mayoría de las leguminosas noduladas examinadas por Schubert y Evans mostraron una eficiencia relativa entre 0,52 y 0,70, excepto algunas especies del grupo "cowpea" (*Vigna unguiculata*, entre otras) donde aparecía una ER próxima a la unidad. Las especies no leguminosas, pero capaces de formar nódulos con actinomicetos (*Frankia*), tales como *Alnus, Myrica* y *Casuarina,* presentaban unos valores de ER también cercanos a la unidad.

La pregunta que debe hacerse, después de conocidos estos hechos, es si de la presencia de actividad hidrogenasa, que permita el reciclaje del H_2 producido por la nitrogenasa, se deriva un incremento de la fijación de nitrógeno o se favorece el crecimiento de las plantas. En este sentido, los datos obtenidos por H. J. Evans, N. Jennings y J. Hanus, en el ya citado Departamento de Botánica y Fitopatología de la Universidad de Oregón, y T. Ruiz-Argüeso, en la Escuela Técnica Superior de Ingenieros Agrónomos de la Universidad Politécnica de Madrid, muestran que, comparando leguminosas del grupo "cowpea" noduladas, unas donde hay liberación de H_2 y otras no, el peso seco y el N_2 total fijado eran un 11 y un 14 por ciento superior respectivamente, en la asociación donde no hay evolución de H_2. En el caso de *R. japonicum*-soja se obtienen incrementos del 24 y del 31 por ciento en el peso seco y en el total de N_2 fijado a favor de la raza que no libera H_2.

De acuerdo con los datos proporcionados por K. R. Carter, N. T. Jennings, F. J. Hanus y H. J. Evans, la información genética para la síntesis de la hidrogenasa capaz de reciclar el H_2 producido por la nitrogenasa parece residir en la bacteria, como se deduce de los experimentos realizados al inocular diferentes cultivos de soja con la misma raza de *R. japonicum*. Estos autores y E. R. Campbell, del departamento de microbiología de la Universidad de Manitoba, han desarrollado un método, basado en la reducción del cloruro de trifeniltetrazolio, en un medio con agar, para seleccionar razas de *R. japonicum* capaces de sintetizar hidrogenasa en los nódulos de soja. Todavía no está bien establecido el hecho de que la reducción del colorante se deba directa o indirectamente a la actividad hidrogenasa. Sin embargo, el uso de este colorante, que es relativamente poco tóxico para *Rhizobium*, ha permitido aislar razas que no presentan actividad hidrogenasa, de gran utilidad en la caracterización bioquímica de este complejo enzimático.

La capacidad para reciclar el H_2 producido vía nitrogenasa es uno de los factores que contribuyen al incremento de la fijación de nitrógeno por los sistemas fijadores. Por tanto, deben hacerse los esfuerzos necesarios para identificar y desarrollar las combinaciones de razas de *Rhizobium* y cultivares de leguminosas que rindan más desde el punto de vista energético, y para incorporar la capacidad de conservación de la energía a los sistemas simbióticos fijadores de nitrógeno, que posean otras características adecuadas desde la óptica de la fijación.

La obtención de mutantes por tratamientos físicos, químicos o bioquímicos (utilización de transposones) y la puesta a punto de una técnica adecuada para la selección de las razas idóneas habrá de permitir incrementar la eficiencia de la fijación. La actuación paralela sobre la liberación de hidrógeno y la fotorrespiración llevará consigo el máximo aprovechamiento de la energía utilizable en la fijación. El trabajo de genetistas de plantas que obtengan, además, variedades de leguminosas más susceptibles a la infección por *Rhizobium* y capaces de establecer una extensa e intensa simbiosis mutualística, contribuirá a solucionar los problemas planteados por la falta de disponibilidad de proteína suficiente.



EN LOS COMIENZOS DEL CULTIVO la actividad fijadora, medida como reducción del acetileno a etileno, presenta una evolución paralela al crecimiento de la planta, expresado en peso seco. Cuando llega la floración, y comienzo de la fructificación, la fijación decae rápidamente como consecuencia de la falta de fotosintato disponible, que es transportado al fruto en formación. Una actuación adecuada en esta etapa que permita prolongar la función nodular puede dar lugar a un rendimiento mayor de la cosecha.

Ciencia y sociedad

Nobel de física y espectroscopía

El premio Nobel de Física de 1981 se ha otorgado a Nicolaas Bloembergen y Arthur Schawlow, por sus aportaciones al desarrollo de la espectroscopía con láser, y a Kay Siegbahn, por sus contribuciones al desarrollo de la espectroscopía de electrones de alta resolución.

El galardón puede considerarse, en cierto modo, como un reconocimiento de la enorme labor realizada en los últimos años en el campo de la investigación sobre la estructura atómica y molecular. El desarrollo de los sistemas láser ha permitido notables avances en espectrometría. Se ha mejorado la resolución con la que se pueden medir las líneas espectrales y la precisión con la que se pueden determinar constantes fundamentales, como son la velocidad de la luz y la constante de Rydberg. A. Schawlow y N. Bloembergen realizaron trabajos básicos para la invención del láser, pero su labor no terminó allí, continuaron su investigación en campos distintos. Bloembergen en la óptica no lineal, Schawlow en el desarrollo de técnicas de alta resolución; pero del trabajo de ambos ha resultado que el láser se haya convertido en un instrumento de investigación inapreciable en espectrometría. Kay Siegbahn, que ha hecho notables aportaciones en el campo de la física nuclear experimental, y cuyas obras son clásicas en dicho sector, ha desarrollado dentro de la física atómica una técnica analítica que utiliza la espectrometría de electrones. Dicha técnica está basada en la determinación precisa de la energía cinética de los fotoelectrones emitidos por una muestra irradiada con rayos X y en la medida de los desplazamientos del espectro energético de los fotoelectrones debido al enlace químico.

Nicolaas Bloembergen

Nació el 11 de marzo de 1920 en Dordrecht (Holanda). Desde 1958 posee la nacionalidad estadounidense. Enseña en la Universidad de Harvard. Especialista en óptica no lineal y resonancia magnética, nuclear y electrónica, realizó trabajos pioneros en los máseres de tres niveles.

Las altas densidades de energía en la zona visible que proporcionan los láseres han permitido la observación, relativamente sencilla, de fenómenos considerados antes meras curiosidades teóricas. Nos referimos a fenómenos de óptica no lineal, cuya explicación tiene que ver con el modo de propagación de la luz a través de un medio sólido. El medio está formado por átomos cuyo núcleo y electrones corticales forman dipolos eléctricos. La radiación electromagnética interacciona con estos dipolos, los hace oscilar y éstos, a su vez, actúan como fuentes de radiación electromagnética. Cuando la amplitud de la vibración de los dipolos es pequeña, la radiación que emiten, en virtud de su oscilación, está relacionada linealmente con la de la radiación incidente. Pero cuando la intensidad de la radiación aumenta, también lo hace la amplitud de la vibración de los dipolos y aparecen no linealidades. Un resultado inmediato es que surjan armónicos de la frecuencia principal de oscilación. Las no linealidades pueden ocurrir en un medio pasivo, tal como el descrito anteriormente, es decir, en un medio donde sus frecuencias características no desempeñan ningún papel; consideraremos, por sus aplicaciones a la espectrometría, los fenómenos de generación óptica de armónicos y de mezcla de frecuencias.

Aunque también las no linealidades se pueden producir en un medio activo donde aparezcan las frecuencias propias. Ejemplos de interés de este caso son las dispersiones estimuladas Raman, Rayleigh y Brillouin. En la ilustración de la página opuesta se muestran tres efectos no lineales. En primer lugar, el de obtención de frecuencia doble a la del láser por un cristal, que es de gran utilidad para convertir luz láser visible en ultravioleta. En segundo lugar, la amplificación paramétrica que permite obtener por mezcla no lineal en un cristal la frecuencia diferencia entre dos dadas y tiene aplicación para obtener fuentes de luz intensas en la parte infrarroja del espectro. En tercer lugar, el efecto Raman estimulado, que permite obtener frecuencias superiores a la frecuencia del láser por emisión de líneas anti-Stokes (en honor de George Gabriel Stokes, físico inglés).

Entre los cristales más utilizados en la generación de frecuencias dobles están el de fosfato dihidrogenado de potasio (KDP) y el de fosfato dihidrogenado de amonio (ADP). Para que la eficiencia de producción de luz de frecuencia doble, (2ω), sea lo mayor posible, la velocidad de propagación de los dos haces, el de la frecuencia original y el de la frecuencia doble, debe ser igual: el índice de refracción para la luz incidente debe ser igual al índice para la frecuencia doble. Esto se puede conseguir fácilmente, si, siguiendo la terminología de la óptica en medios cristalinos, la luz fundamental es luz "ordinaria" polarizada linealmente y el segundo armónico es una onda "extraordinaria". Si el ángulo que forma la dirección de la luz con respecto al eje óptico es $\theta \simeq 55^{\circ}$ para el cristal KDP, tal como se muestra en la figura, $n_o(\omega) = n_e(2\omega)$. El ángulo de orientación óptimo es función de la longitud de onda y ha de ajustarse para cada una de ellas. De este modo se puede conseguir un rendimiento en la conversión de frecuencia de hasta el 20 por ciento.

Vale la pena hacer un breve comentario sobre los fenómenos multifotónicos, que ofrecen grandes perspectivas de aplicación, tanto científica como industrial. En la ilustración de la página 38 se muestra, en la parte *a*, el esquema de un experimento de ionización multifotónica de átomos, empleando láseres de gran potencia. Los láseres que se emplean en la actualidad para tal fin son los de rubí y neodimio. Como la energía de los fotones emitidos por estos láseres es de 1,8 y 1,2 electronvolt, respectivamente, y las energías de ionización de los átomos se encuentran, normalmente, en el margen de 5 a 15 electronvolt, se comprende que haga falta suponer la suma de los efectos de un gran número de fotones para explicar la ionización. Como la probabilidad de ionización es proporcional (si no existen resonancias intermedias) a la potencia n de la intensidad del láser (donde n es el número de fotones necesario para ionizar el átomo), una representación gráfica logarítmica de la corriente de ionización en función de la intensidad del láser (I_L) será una recta, cuya pendiente valdrá el número de fotones necesarios para la ionización.

En el caso de las moléculas se puede conseguir la disociación multifotónica mediante la acción de potentes láseres de CO_2, cuya energía por fotón es de 0,12 electronvolt. El número de fotones necesarios es del orden de 30. Una característica muy notable de este fenómeno es que, variando la longitud de onda de la radiación excitadora, se puede conseguir disociar selectivamente moléculas que contengan diferentes isótopos. Esto se logra excitando en condiciones de resonancia los primeros niveles vibracionales que presentan distinta separación según su composición isotópica.

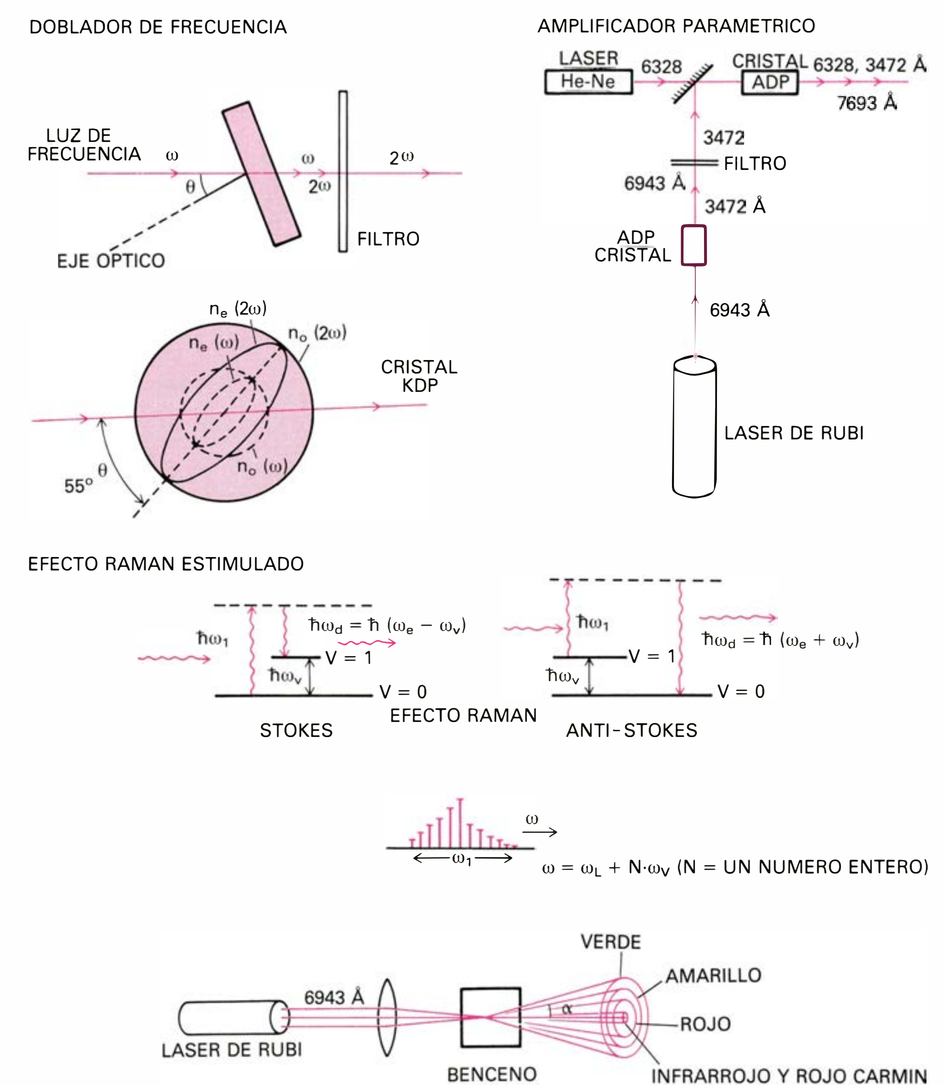
Arthur L. Shawlow

Arthur L. Schawlow nació el 5 de mayo de 1921 en Mount Vernon, en el estado de Nueva York. Se graduó en la Universidad de Toronto (Canadá). Es profesor de la Universidad de Stanford desde 1961. Junto con su cuñado Charles H. Townes estudió la posibilidad del máser óptico, lo que denominamos actualmente láser. C. H. Townes compartió el premio Nobel de física en 1964 con N. Bassov y A. Prokhorov por sus investigaciones fundamentales de radiofísica cuántica que condujeron a la creación de máseres y láseres. A. L. Schawlow continuó sus estudios aplicando el láser a la espectroscopía de alta resolución. Entre sus creaciones están la espectrometría de absorción saturada, la eliminación de la anchura Doppler en transiciones multifotónicas, la resolución de las componentes de la estructura fina de la línea $H\alpha$ del hidrógeno por absorción saturada y la aplicación de la espectrometría de polarización al estudio de espectros moleculares y atómicos [véase "El espectro del hidrógeno atómico", por Theodor W. Hänsch, Arthur L. Schawlow y George W. Series; INVESTIGACIÓN Y CIENCIA, mayo de 1979].

En lo que sigue explicaremos brevemente algunas de estas realizaciones, señalando su interés en el estudio de átomos y moléculas. En la sección superior de la ilustración de la página 39 se muestran algunos métodos y resultados típicos de la espectrometría de saturación. El haz de luz de un láser de colorante, de longitud de onda variable, se divide en dos partes, un haz intenso que ha de saturar la absorción del medio en el que se propague y un haz mucho menos intenso de prueba. A los dos haces se les hace atravesar el medio, en el experimento de Schawlow era una descarga de H, en direcciones opuestas a fin de eliminar el ensanchamiento Doppler. El haz intenso selecciona, saturando su absorción, aquellos átomos que tienen velocidad longitudinal igual a cero. El haz de prueba no es absorbido cuando interacciona con este tipo de átomos. De este modo se obtiene el espectro que constituye la línea $H\alpha$ en donde se pueden observar las líneas de la estructura fina, incluido el desplazamiento Lamb. Algunos fenómenos espurios como las denominadas resonancias de cruce, que aparecen cuando dos transiciones comparten el mismo nivel superior o inferior, son fácilmente detectables.

Por este método se ha obtenido para la constante de Rydberg el valor $10.973.731{,}513 \pm 0{,}085\ m^{-1}$, siendo por tanto una de las constantes fundamentales mejor conocidas; se ha podido determinar el desplazamiento Lamb, por métodos ópticos, con una precisión superior al 1 por 1000.

En los laboratorios en los que se realiza espectrometría con láser existen muchos dispositivos para espectrometría por absorción saturada basados en los principios anteriores. La parte *B* de la misma ilustración recoge otro método utilizado en experiencias de fluorescencia de haces atómicos. A un haz de

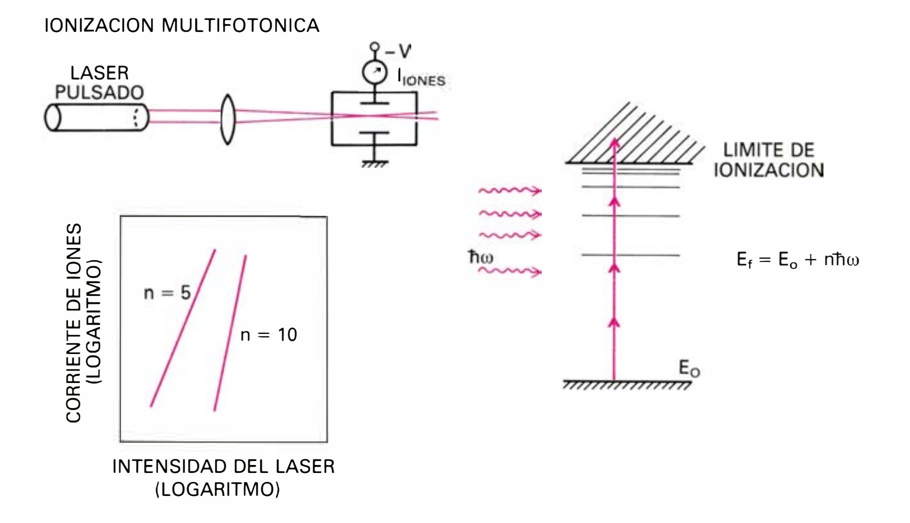


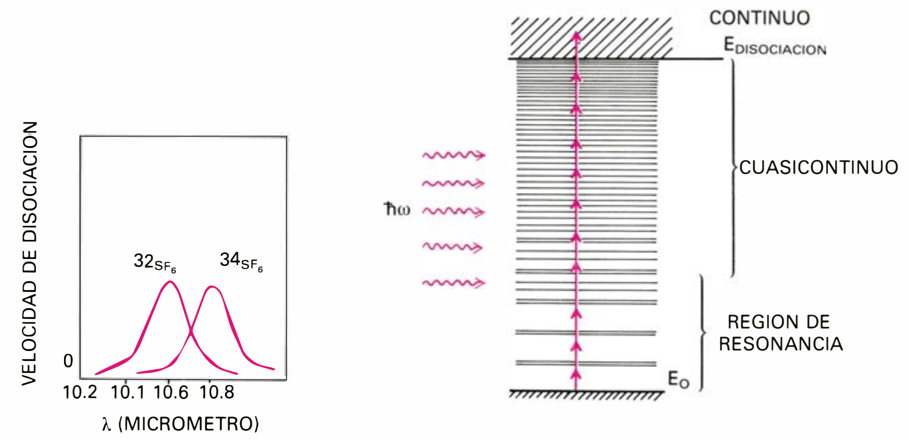
Algunos fenómenos no lineales útiles para obtención de luz de diferente longitud de onda a partir de la del láser original. **A**. *Obtención de la frecuencia doble mediante un cristal KDP. Obsérvese que el eje óptico se sitúa con respecto al de incidencia de la luz de modo que n_o (ω) = n_e (2ω).* **B**. *Ejemplo de oscilador paramétrico para extensión de la longitud de onda de un láser hacia longitudes de onda más largas. La frecuencia de la luz de 7693 Å es la diferencia de la de 6328 Å y de 3472 Å.* **C**. *Efecto Raman estimulado. En la parte superior se ilustran la generación de líneas Stokes y anti-Stokes en dicho efecto. En la parte inferior, el resultado de un experimento clásico. Los anillos de longitud de onda más corta que se observan, así como su apertura angular, fueron explicados teóricamente por N. Bloembergen.*

láser se le hace atravesar un haz de iones y volver por el mismo camino tras sufrir reflexión en un espejo. El correspondiente perfil de la línea de fluorescencia muestra una depresión muy aguda correspondiente a la saturación de los átomos cuya componente transversal de la velocidad en el haz es nula.

A. L. Schawlow obtuvo el espectro de la línea $H\alpha$ mediante otro método, llamado espectrometría de polarización, basado en que los cambios en la polarización de la luz se pueden detectar con más precisión que los cambios en la intensidad. El dispositivo experimental que, a este respecto, se muestra en la ilustración de la página siguiente es semejante al de la espectrometría de saturación. Al haz saturante se le polariza circularmente. La intensidad del haz de prueba se encuentra anulada a su salida porque ha de atravesar dos polarizadores lineales cruzados, y sólo puede pasar luz si el haz saturante ha interaccionado con el medio, de tal modo que éste haga girar el plano de polarización del haz de prueba. Al igual que en el caso de la espectrometría de saturación, el efecto de aumento de intensidad de salida del haz de prueba sólo se producirá si los dos haces interaccionan con los átomos de velocidad longitudinal nula en un cierto instante, eliminándose el efecto Doppler.

Otro modo de anular el efecto Doppler consiste en someter a los átomos a la acción simultánea de dos haces de direcciones opuestas y conseguir transiciones por dos fotones. En el sistema de referencia del átomo, este verá ambos fotones con desplazamientos Doppler iguales y de signo contrario, anulándose éstos en la transición por dos fotones, independientemente del valor que tenga la velocidad del átomo. En resumen, los elaborados métodos desarrollados por A. L. Schawlow permiten realizar determinaciones espectrométricas con una precisión en la resolución que hubiera sido impensable hace 20 años.





A. *Ionización multifotónica de un átomo. La representación gráfica logarítmica de la corriente de iones en función de la intensidad del láser es, en ausencia de resonancias, una recta de pendiente* **n**, *donde* **n** *es el número de fotones necesarios para ionizar el átomo.* **B**. *Disociación de una molécula por acción de un láser de infrarojo, tal como el de CO_2. En los niveles de la molécula se observa una región de niveles discretos en la que se puede conseguir una resonancia selectiva con respecto a la composición isotópica de la molécula y una región de cuasicontinuo, no selectiva. A la derecha se muestran los resultados obtenidos por el SF_6.*

Kay Siegbahn

Nació el 20 de abril de 1918 en la localidad de Lund, en el sur de Suecia. Es director del Instituto de Física de la Universidad de Uppsala. Su padre obtuvo también el premio Nobel de Física en 1924; realizó importantes estudios sobre rayos X. Kay Siegbahn debe su fama a los trabajos publicados sobre espectrometría de electrones y la creación de la técnica denominada ESCA (Electron Spectroscopy for Chemical Analysis), amén de sus muchas aportaciones en el campo de la física nuclear y de las técnicas experimentales relacionadas.

Como dijimos antes, ha recibido el galardón por su aportación a la espectrometría de electrones. Si detenemos un momento nuestra mirada en la ilustración superior de la página 40 veremos un esquema aproximado de los márgenes de energías de los electrones cubiertos por determinados fenómenos atómicos o moleculares. El primero que está indicado es el margen de la energía de enlace de los electrones en átomos y moléculas (designado por *e* en la figura). Según la capa de los electrones, la energía de enlace puede cifrarse entre unos pocos electronvolt (eV), para los electrones más externos, hasta ser de cerca del centenar de kiloelectronvolt (keV) para las capas más profundas de los elementos pesados. Los desplazamientos en la energía de ligadura debidos al enlace químico (designados por *q*) se encuentran en el margen aproximado de 0,1 a 10 eV. Se señala luego el margen de anchura de los niveles (*n*) atómicos internos, que está relacionada con la vida media del nivel, siendo igual al cociente entre la constante de Planck dividida entre 2π y la vida media. A continuación viene (*A*), el margen de energías de los electrones Auger, que son los emitidos por un átomo cuando, producido un hueco, el exceso de energía no se emite como energía electromagnética, sino que se comunica a otro electrón de la corteza que es emitido con una cierta energía

cinética. Por último se indica el margen de energías cinéticas que pueden poseer los electrones producidos por autoionización (*au*, en la figura), proceso que tiene lugar cuando en un átomo o molécula se excita un electrón hasta niveles de energía que están por encima del límite de ionización, produciéndose la ionización espontánea de dicho electrón.

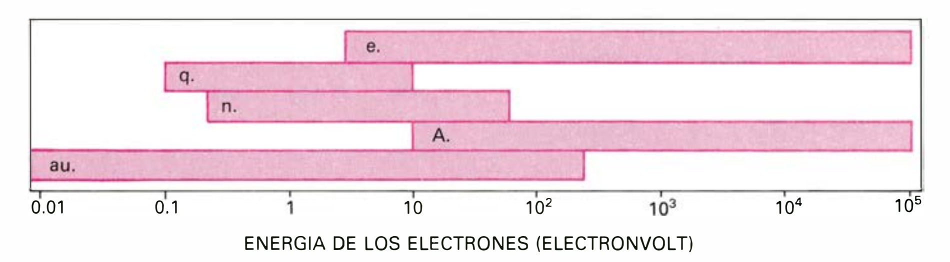
Los procesos indicados, aunque no constituyen una enumeración completa, nos proporcionan una idea de la información que puede obtenerse sobre la estructura atómica y molecular mediante la determinación precisa de energías de los electrones emitidos cuando se excita la corteza por radiación electromagnética o impacto de partículas. Utilizando el efecto fotoeléctrico se puede deducir la energía de enlace de un electrón como diferencia entre la energía del fotón incidente y la cinética del electrón emitido. La información obtenida en los primeros experimentos estuvo limitada por la mala resolución del espectro de energías de los fotoelectrones, consecuencia a su vez de la poca resolución de los espectrómetros usados y de la poca monocromaticidad de las fuentes de rayos X utilizadas. Después de la segunda guerra mundial se consiguió mejor resolución, lo que proporcionó información adicional de gran valor para la química; Siegbahn, en la década de los sesenta, lo asentaría. Por tanto, se puede afirmar que entre los métodos instrumentales de la química, la espectroscopía de electrones es relativamente nuevo. La resolución espectral conseguida últimamente lo ha sido a través de dos líneas de desarrollo diferentes, una mediante nuevo diseño de espectrómetros y otra monocromatizando la fuente de rayos X.

K. Siegbahn obtuvo, en 1957, el primer espectro de alta resolución usando un espectrómetro magnético sin hierro, al que se había añadido un tubo de rayos X que emitía las líneas $K\alpha_1$, y $K\alpha_2$ del molibdeno (17,4 keV). El aparato se usó para obtener las energías de enlace de los electrones internos del cobre. Al año siguiente, se observó un desplazamiento de 4,4 eV entre las energías de enlace de los electrones 1*s* y 2*s* en cobre y en óxido cúprico. Tras un detenido estudio de las energías de enlace de 232 niveles en 76 elementos, Siegbahn y sus colaboradores publicaron, en 1964, un artículo titulado *Espectroscopía de electrones para el análisis químico*, señalando que los desplazamientos químicos se observaban

Espectroscopía láser de alta resolución. **A**. *Espectroscopía de saturación. Obsérvese el perfil Doppler de la línea Hα (de trazos) y el resultado obtenido con el dispositivo de la figura superior. En la parte inferior se muestra el mismo principio aplicado a la luz emitida en la fluorescencia atómica en un experimento con haces iónicos. Se encuentra aplicado a la medida de desplazamientos isotópicos.* **B**. *Espectroscopía de polarización. P_1 designa el polarizador circular, P_2 y P_3, los polarizadores lineales cruzados. Las resonancias espúrias de cruce aparecen de forma distinta, pero en el mismo lugar, que en el método de saturación.* **C**. *Espectroscopía por dos fotones. La figura muestra la excitación de la transición "prohibida" 1s-2s.*

siempre, y cabía relacionarlos con el estado de valencia del átomo cuyo electrón interno había sido expulsado, y que las intensidades de los máximos observados podían relacionarse con el número de átomos presentes con un determinado tipo de enlace.

Los rayos X se generan por impacto de un haz de electrones de alta energía sobre un blanco metálico que emite un electrón interno, rellenándose el hueco mediante una transición entre los electrones de la capa interna con emisión de fotones. Pero su espectro no es monocromático y se excitan varias líneas, junto con un fondo continuo de radiación de frenado (*Bremstrahlung*). La intensidad relativa varía considerablemente con el átomo que sirve de blanco; mas, para los materiales utilizados en la espectrometría de fotoelectrones, que son el aluminio y el magnesio, las

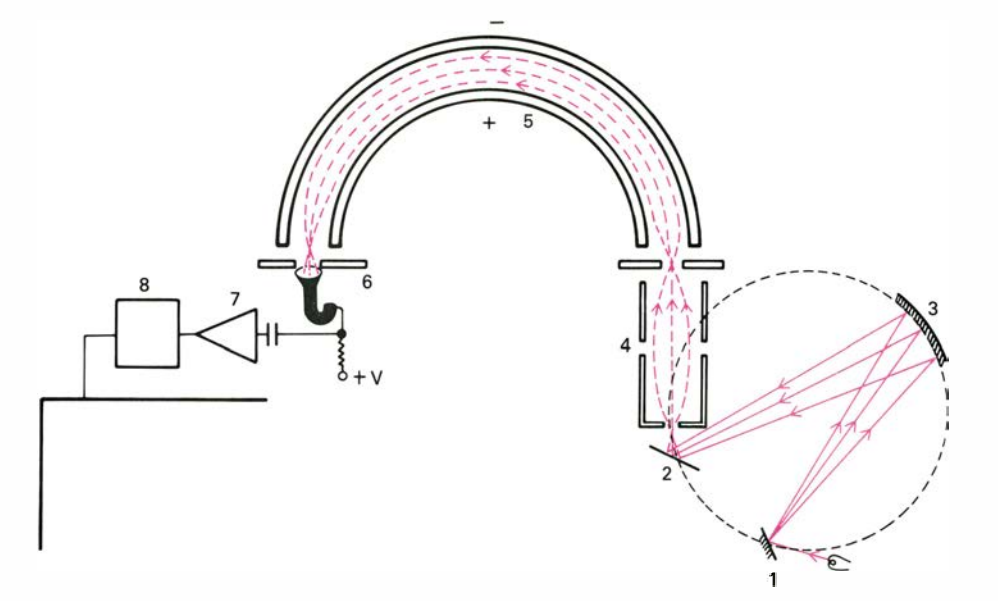


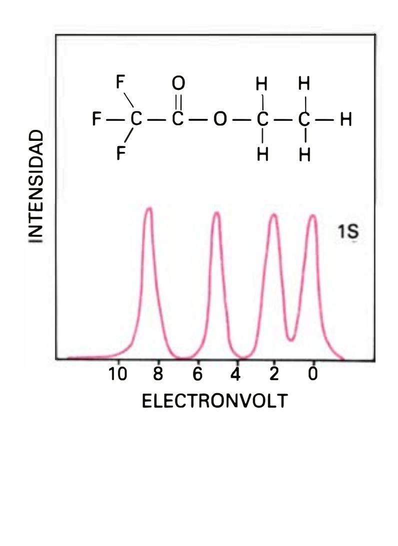
Márgenes de energías de los electrones en átomos o moléculas; (e) energía de ligadura; (q) desplazamientos de la energía de ligadura debidos al tipo de enlace químico; (n) anchura de los niveles energéticos internos; (A) margen de energías cinéticas de los electrones emitidos por efecto Auger; (au) energías cinéticas de electrones producidos en la autoionización de átomos o moléculas.

únicas líneas son las $K\alpha_1$ y $K\alpha_2$, que se solapan, incrementando la anchura de la radiación X. Esta anchura se reduce monocromatizando la radiación. Siegbahn incorporó un monocromador de rayos X a los espectrómetros de fotoelectrones. La radiación X se dispersa por reflexión de Bragg de un cristal curvado que actúa como red de difracción. La fuente de rayos X, el cristal y la muestra a excitar se encuentran en el círculo de Rowland correspondiente. La resolución de un espectrómetro de este tipo puede ser de 0,2 eV. El dispositivo produce, no obstante, una apreciable pérdida de intensidad, que se puede reducir colocando la muestra en el círculo de Rowland, de tal modo que cada punto de su superficie está expuesto a radiación monocromática cuya longitud de onda varía a lo largo de la muestra, y el analizador de electrones se dispone de tal modo que compense este efecto. El uso de rayos X monocromatizados no sólo produce líneas estrechas, sino que elimina completamente los efectos debidos a líneas X satélites y a radiación de espectro continuo. Para poderla usar se han de emplear, empero, espectrómetros de electrones de diseño muy elaborado.

Los espectrómetros usados actualmente en trabajos de alta resolución son del tipo electrostático. En la ilustración inferior está representada una sección de un espectrómetro electrostático hemiesférico que posee notables propiedades de enfoque y luminosidad. La selección de energías se realiza variando la diferencia de potencial entre las dos semiesferas concéntricas. También se utiliza ampliamente el selector de energía cilíndrico de 127°. Para la detección de los electrones seleccionados en energía se utilizan multiplicadores por emisión secundaria de canal continuo y, actualmente, detectores múltiples que permiten la detección simultánea de electrones de energías próximas, lo que aumenta el rendimiento del sistema.

Los desplazamientos químicos se pueden observar en los electrones internos de casi todos los elementos. Significa ello que la espectrometría de fotoelectrones procedentes de la zona interna constituye una técnica muy valiosa para investigar el enlace químico. Por otra parte, como las energías de enlace son características del tipo de átomo, los constituyentes atómicos de un compuesto se pueden determinar directamente; además, el área de los picos corresponde al número de átomos presentes. Por ejemplo, en la citada ilustración, se muestra el resultado obtenido para el etiltrifluoroacetato en fase gaseosa. Se observan cuatro máximos correspondientes a la energía de enlace de electrones de carbono en el orbital 1*s* (energía de enlace ≃ 290 eV). Cada máximo corresponde a un átomo de carbono en diferentes condiciones de enlace. Para finalizar, debemos indicar que otra rama de la espectrometría de fotoelectrones es su aplicación al estudio de superficies en estado sólido, campo en el que está proporcionando resultados de gran interés.





Un dispositivo típico para el estudio de energía de electrones producidos por efecto fotoeléctrico en una molécula: **1***) blanco que produce rayos X por impacto de electrones;* **2***) muestra excitada por rayos X monocromatizados que produce fotoelectrones;* **3***) monocromatizador de rayos X de cristal curvo;* **4***) sistema de lentes electrostáticas de enfoque;* **5***) espectrómetro de electrones hemisférico representado en sección;* **6***) detector de electrones (multiplicador por emisión secundaria de canal continuo);* **7***) amplificador de impulsos y* **8***) frecuencímetro analógico. En la parte inferior izquierda se representa un espectro de fotoelectrones procedentes del nivel 1s del carbono en el compuesto etiltrifluoroacetato. Los cuatro máximos se encuentran desplazados al estar los correspondientes átomos de carbono con un enlace químico distinto.*

Júpiter y Saturno

Dos modelos rivales tratan de describir los compañeros gigantes del Sol. En un modelo, los vientos están confinados a una delgada capa junto a la superficie; en el otro, los vientos se extienden por las profundidades fluidas de los planetas

Andrew P. Ingersoll

Hubo muchos momentos emocionantes durante el paso de los vehículos espaciales Voyager por la proximidad de Júpiter en 1979 y de Saturno en 1980 y 1981; mas, para mí, los más memorables se produjeron al principio de la primera pasada, cuando vimos a cámara rápida las imágenes de las turbulentas nubes de Júpiter hechas por el *Voyager 1* en su larga aproximación al planeta. La secuencia comprimía una historia de 30 días del tiempo meteorológico de Júpiter en una película de un minuto de duración. Como científico que estudia las atmósferas de los planetas, estaba familiarizado con los cinturones castaños y las zonas blancas de Júpiter: bandas coloreadas de nubes de miles de kilómetros de anchura que rodean el planeta a latitud constante. Y había llegado a aceptar la hipótesis de que la Gran Mancha Roja de Júpiter era una borrasca tan grande como la Tierra y que persistía a lo largo de siglos.

Pero no me esperaba los intrincados movimientos que reveló la película en cámara rápida. La secuencia estaba sincronizada con la rotación de Júpiter, de modo que la Gran Mancha Roja parecía estacionaria, con la atmósfera adyacente arremolinándose a su alrededor. Surgían detalles brillantes y pequeños que se esfumaban a los pocos días. Motas que incidían sobre la Mancha Roja procedentes del Este parecían dejarse arrastrar por su circulación antihoraria, rodeándola en una o dos semanas y dividiéndose después en dos partes, de las que una se soldaba a la Mancha y la otra regresaba hacia el Este. En otros lugares se estaban formando manchas, uniéndose y dividiéndose a intervalos de días. A escalas inferiores a los 1000 kilómetros, aquello era el caos. ¿Cómo podían las estructuras mayores perdurar y retener sus coloraciones distintivas en una atmósfera tan bien mezclada? Nos hallábamos ante una incógnita más difícil de resolver que lo fuera antes de la llegada de los vehículos espaciales.

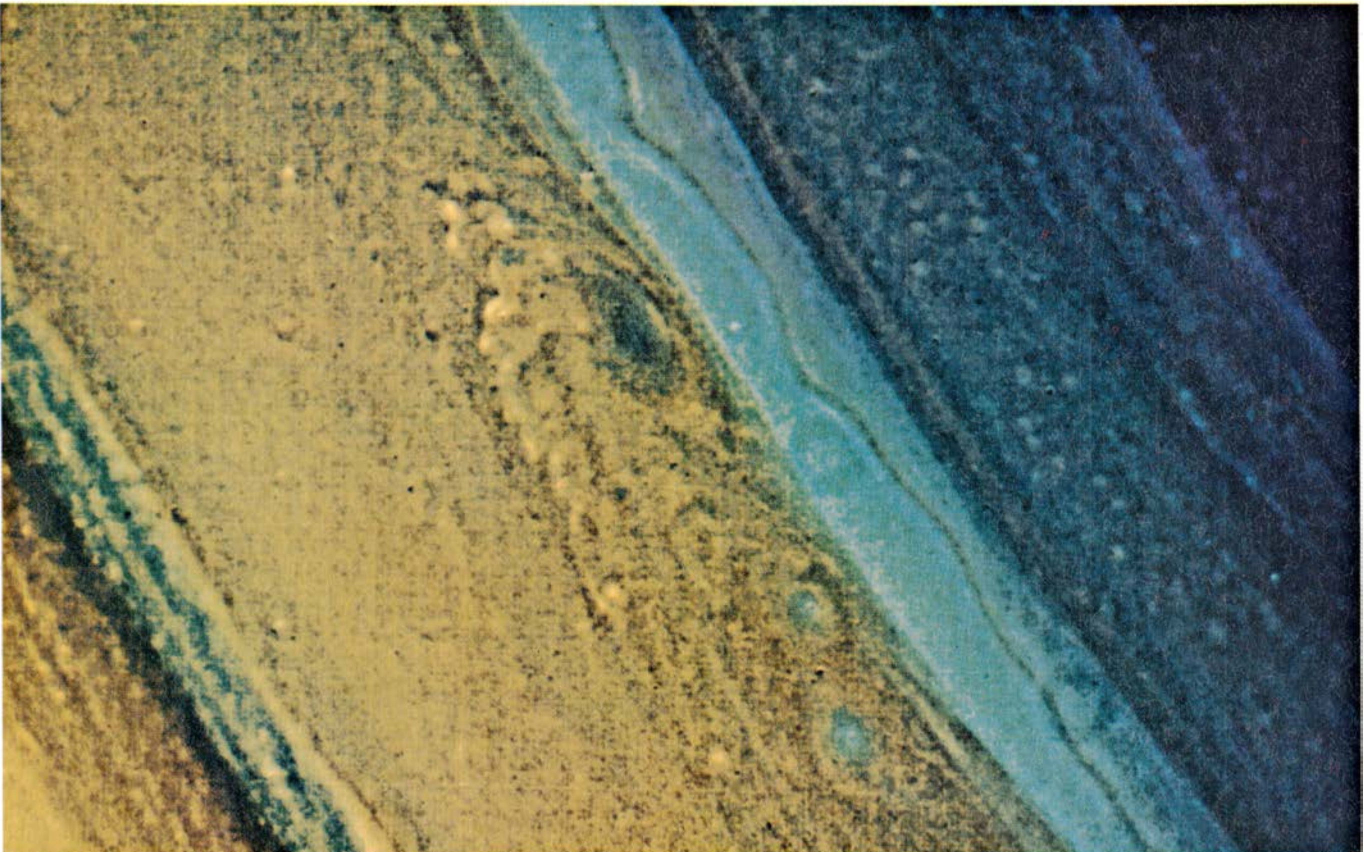
Y sigue siendo una cuestión pendiente. A pesar de todo, el análisis de las imágenes hechas por el *Voyager 1* y el *Voyager 2* revela orden en el caos. Así, muestra que muchos aspectos de la circulación atmosférica de Júpiter y Saturno recuerdan configuraciones de la atmósfera y océanos de la Tierra. Se invocan ahora cierto número de modelos teóricos y experimentos de laboratorio para explicar dicha circulación. Modelos que difieren en las hipótesis que formulan acerca de una cuestión básica no resuelta. ¿Están los movimientos de la atmósfera de Júpiter y Saturno confinados a una delgada capa iluminada, de 100 a 200 kilómetros de espesor, donde se forman las nubes y la presión atmosférica es sólo algunas veces mayor que la presión al nivel del mar en la Tierra? ¿O se extienden los movimientos a lo largo de decenas de miles de kilómetros hasta la cima de una zona de hidrógeno metálico, donde la presión es tres millones de veces superior a la presión al nivel del mar en la Tierra?

Masa, densidad y estructura interna

Al abordar estas preguntas, me ceñiré a Júpiter y Saturno en sí mismos, excluyendo por tanto satélites y anillos. Comenzaré por las propiedades planetarias en gran escala: la masa, densidad y composición. Ello nos llevará al examen de la estructura interna de cada planeta y del origen del calor interno que radia hacia el espacio. Describiré, a continuación, la estructura y la química de sus atmósferas. Pasaré revista, por último, a las distintas explicaciones que se han avanzado sobre la dinámica atmosférica. De lo que se trata aquí es de averiguar por qué las características generales del flujo atmosférico persisten a lo largo de decenios o siglos en Júpiter y Saturno y apenas duran días o semanas en la Tierra.

No existe teoría alguna que explique, de forma convincente, la masa de los dos principales compañeros del Sol. La de Júpiter es de 318 veces la masa de la Tierra, lo que representa alrededor de una milésima parte de la solar. La masa de Saturno supera en 95 veces la de la Tierra. Presumiblemente, esos valores reflejan la cantidad de materia abandonada en órbita alrededor del Sol poco después de que éste se formara. Ninguna teoría explica tampoco de forma convincente las abundancias relativas de los diversos elementos químicos de que constan Júpiter y Saturno; aunque sí se parecen a las halladas en el Sol, Júpiter y Saturno son los únicos planetas del sistema solar que están formados predominantemente de hidrógeno y helio. No habría otras substancias que pudieran dar a Júpiter su densidad, de sólo 1,33 gramos por centímetro cúbico, a las presiones y temperaturas que caracterizan al planeta. La densidad de Saturno es todavía menor, de 0,69 gramos por centímetro cúbico, porque su menor masa comporta un grado menor de autocompresión gravitatoria.

Mercurio, Venus, la Tierra y Marte están constituidos de materia más densa: propiedad ésta que varía desde 3,9 hasta 5,5 gramos por centímetro cúbico, es decir, algunas veces la densidad de Saturno. En general, están compuestos por rocas: los metales más abundantes y sus óxidos. Las densidades de Urano y Neptuno doblan la de Saturno, aunque su autocompresión es menor. Sus elementos constituyentes más probables son, por tanto, oxígeno, carbono y nitrógeno: el tercero, cuarto y quinto entre los elementos más abundantes en el sistema solar. (Los dos primeros son el hidrógeno y el helio.) A las temperaturas que caracterizan el sistema solar exterior, el oxígeno, carbono y nitrógeno deben combinarse con el hidrógeno disponible para formar agua, metano (CH_4) y amoniaco (NH_3). En las superficies de Urano y



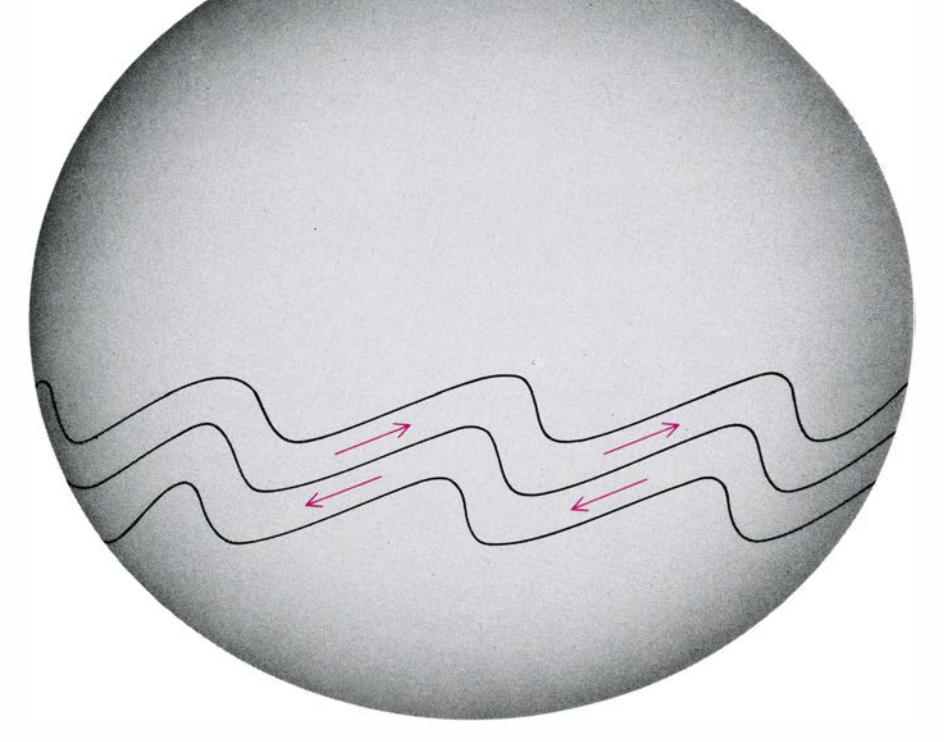
CORRIENTES TURBILLONARIAS en la atmósfera de Júpiter y Saturno. Son desviaciones caóticas con respecto a una corriente que, por lo demás, está formada por cinturones alternados de vientos del Oeste y del Este. Las latitudes medias septentrionales de Júpiter (*arriba*) se fotografiaron en color exagerado por el vehículo espacial *Voyager 1*, el 2 de marzo de 1979. La cinta anaranjada que corta diagonalmente la parte inferior derecha representa una corriente constante de viento del Oeste, cuya velocidad es de unos 130 metros por segundo. Las líneas sinuosas hacia la parte superior representan torbellinos cuyas velocidades son de unos 30 metros por segundo con respecto a las corrientes más laminares. El *Voyager 2* fotografió, el 2 de agosto de 1981, las latitudes medias septentrionales de Saturno en color exagerado (*abajo*). La línea sinuosa dentro de la cinta azul claro es una corriente que se mueve hacia el Este a 150 metros por segundo. El óvalo oscuro y las formaciones con aspecto de humareda debajo de él son remolinos que avanzan hacia el Oeste.

Neptuno, estos compuestos se encuentran en forma sólida ("hielos"); en el interior son líquidos. No hay fácil explicación del enriquecimiento de hielos con respecto al hidrógeno y helio en los planetas exteriores a Saturno. El empobrecimiento de gases y hielos en el sistema solar interno refleja probablemente las altas temperaturas reinantes en la cercanía del Sol en la primitiva fase de la vida del sistema solar.

Conocida la masa y composición de un planeta como Júpiter o Saturno, se puede tratar de inferir su estructura interna. El tamaño y la densidad del planeta se ajustan de suerte tal que la presión exterior de la materia comprimida equilibra exactamente la atracción gravitatoria hacia dentro en cualquier punto del interior del planeta. El resultado es un estado de equilibrio hidrostático. Si el planeta se halla en rotación, hay que sumar una fuerza adicional que interviene en el equilibrio: la fuerza centrífuga, dirigida hacia fuera, que resulta de la rotación. La fuerza hacia fuera de una masa en rotación es proporcional al cuadrado de la distancia de la masa al eje de rotación. De aquí que un planeta en rotación se aplaste: su radio polar es inferior a su radio ecuatorial. El grado de aplastamiento dependerá de la distribución interna de su masa. Por ejemplo, dos planetas de la misma masa y velocidad de rotación diferirán en su grado de aplastamiento si la masa está concentrada cerca del centro en uno y más alejada del centro en el otro. De los dos, el último será el más aplastado. El grado de aplastamiento constituye, pues, una sonda sensible que nos permite desentrañar la estructura interna de un planeta.

Tanto Júpiter como Saturno tienen un período de rotación de unas 10 horas. Ambos planetas están, además, algo aplastados. El radio ecuatorial de Júpiter es un 6,5 por ciento mayor que su radio polar; el de Saturno, 9,6 por ciento mayor. Las medidas del campo gravitatorio de los planetas implican un correspondiente grado de concentración de la masa hacia el plano ecuatorial. La incorporación de las medidas en modelos de la estructura interna de Júpiter y Saturno condujo a William B. Hubbard Jr., de la Universidad de Arizona, y a V. Zharkov y V. Trubitsyn, del Instituto de Física de la Tierra de Moscú, a una conclusión ulterior: la densidad de su núcleo hace imposible que esté formado de hidrógeno y helio comprimidos. La presión dentro de cada planeta es simplemente insuficiente para producir las densidades centrales requeridas a partir de una mezcla de esos dos elementos. Al parecer, Júpiter tiene un núcleo de roca y hielo que constituye aproximadamente el 4 por ciento de su masa, y Saturno tiene un núcleo análogo que constituye alrededor del 25 por ciento de su masa. Cada núcleo puede ser la "semilla" sobre la cual se condensó el resto del planeta a partir de gases, cuando se formó el sistema solar. O tal vez los núcleos se formaron después, a raíz de una redistribución de la materia en el interior.

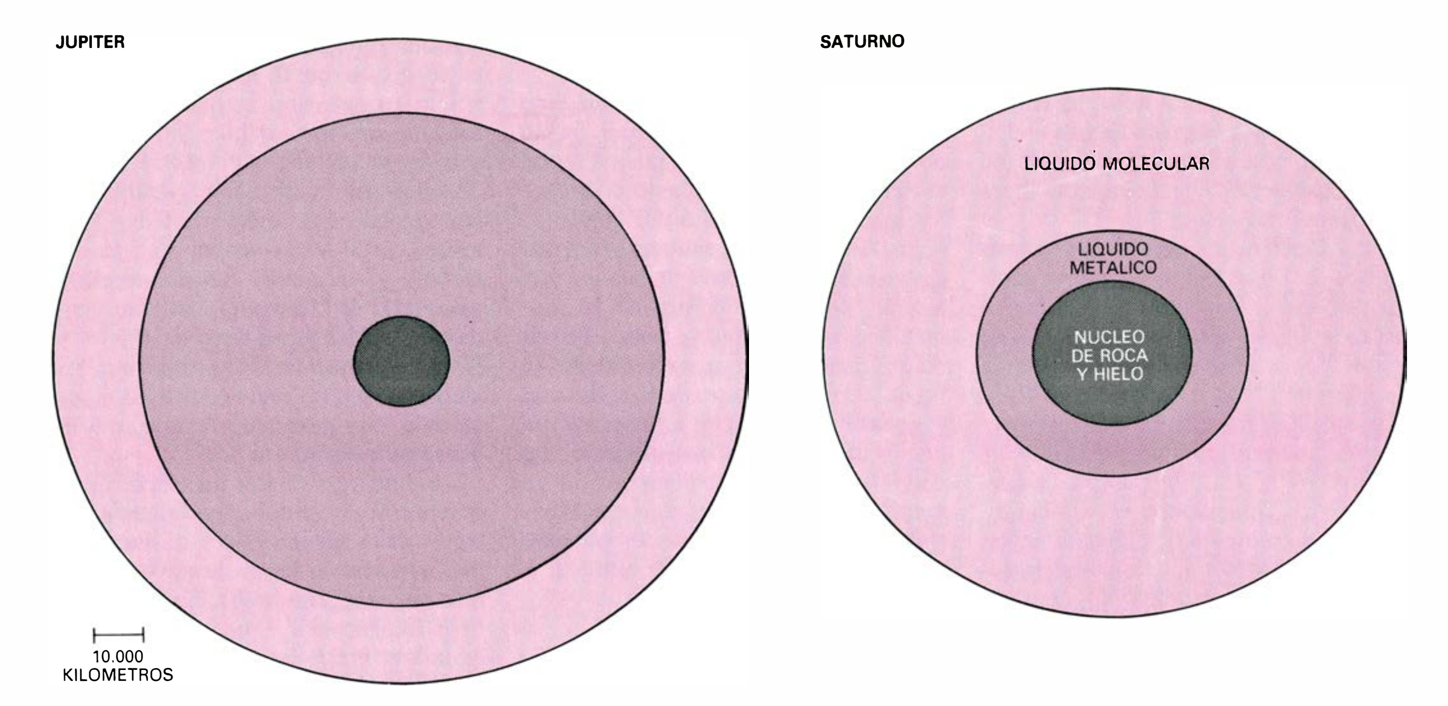
Para conocer mejor la estructura interna de Júpiter y Saturno vale la pena recurrir a la mecánica cuántica, y, dentro de ella, a su descripción del comportamiento de átomos y moléculas cuando están comprimidos. Según el principio de exclusión de la física moderna, los electrones ligados a protones en un conjunto comprimido de moléculas de hidrógeno pueden ocupar el mismo volumen reducido únicamente saltando a niveles más altos de energía. Para cierto grado de compresión (y, por tanto, cierta cantidad de energía) no quedan ya ligados a protones individuales, sino que permanecen libres para errar por el interior de una mezcla eléctricamente neutra de protones y electrones. El hidrógeno se convierte entonces en un metal. Los cálculos realizados por Edwin E. Salpeter, de la Universidad de Cornell, y David J. Stevenson, que actualmente está en el Instituto de Tecnología de California, demuestran que la transición del hidrógeno molecular al hidrógeno metálico se produce a casi la misma presión crítica (tres millones de atmósferas terrestres) tanto en Júpiter como en Saturno. Como el primer planeta tiene mayor masa que el segundo, la presión crítica se alcanza más cerca de la superficie. En Júpiter, la distancia desde el centro del planeta hasta la transición metálico-molecular varía de 0,75 a 0,80 veces la distancia que media entre el centro y la superficie. En Saturno está entre 0,45 y 0,50.



EL INTERCAMBIO DE CANTIDAD DE MOVIMIENTO de los remolinos a los cinturones de vientos del Este y del Oeste fue estudiado en la Tierra por Victor P. Starr y sus colegas del Instituto de Tecnología de Massachusetts; las imágenes de los Voyager sugieren que intercambios semejantes suministran cantidad de movimiento a los vientos del Oeste y del Este en Júpiter y Saturno. Se muestra aquí cómo los remolinos en un modelo hipotético de planeta inyectan cantidad de movimiento del Oeste hacia el Norte y del Este hacia el Sur. Esto aumenta la diferencia de velocidad entre las corrientes medias del Oeste y del Este.

Fuentes de calor interno

Las medidas tomadas por instrumentos montados en los vehículos espaciales Voyager, así como en el Pioneer, que llegó a Júpiter ya en 1973, implican que la potencia que este planeta emite (en forma de radiación infrarroja) vale de 1,5 a 2,0 veces la que absorbe (en forma de radiación solar). La potencia que emite Saturno es de 2,0 a



INTERIOR de Júpiter (*a la izquierda*) y de Saturno (*a la derecha*). En uno y otro planeta del Sistema Solar, dicho interior consta de tres capas. La más externa de ellas es una mezcla líquida de hidrógeno molecular y de helio. La más interna es un núcleo de roca e hielo. En la capa intermedia, una presión superior a tres millones de atmósferas terrestres transforma el hidrógeno en una mezcla líquida de protones y electrones errantes. Por cuya razón, el hidrógeno de la capa intermedia es un metal. La rotación de cada planeta lo aplasta ligeramente hacia los polos. El aplastamiento es mayor en Saturno.

3,0 veces la que absorbe. En consecuencia, tanto Júpiter como Saturno poseen fuentes de calor internas. Sin embargo, ninguno de los dos tiene masa suficiente para que la autocompresión gravitatoria haya iniciado la fusión nuclear; vale decir: no son estrellas. Antes bien, su calor interno debe representar la conversión de la energía potencial gravitatoria que quedó disponible cuando el planeta se contrajo a partir de una nube de gas comenzando hace unos 4600 millones de años. James B. Pollack y sus colegas, del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA), han desarrollado modelos de la historia de los planetas gigantes y llegan a la conclusión de que el interior de Júpiter y Saturno está hoy todavía caliente.

Pero, ¿hasta qué punto? La respuesta se deduce de la ley termodinámica que afirma que el calor pasa de los lugares más calientes a los más fríos. Específicamente, una mezcla de hidrógeno y helio en Júpiter o Saturno (una mezcla metálica incluso) no puede transportar calor desde el centro del planeta hacia la periferia, a menos que exista un apreciable grado de decrecimiento de la temperatura con la distancia al centro. Empero, el grado de decrecimiento está limitado por la convección: la redistribución de un fluido en donde las partículas calientes se elevan y descienden las más frías.

Brevemente, la convección mezcla un fluido hasta que el decrecimiento de su temperatura con la altura se hace igual al coeficiente de enfriamiento por elevación adiabática, o ritmo al cual se enfriará una partícula fluida ascendente cuando no intercambia calor con el medio que la rodea. En lugar de perder energía por intercambio de calor, la partícula pierde energía al efectuar trabajo contra el medio circundante al dilatarse. Sobre este presupuesto, se puede demostrar que el gradiente de temperatura en los interiores de Júpiter y Saturno es próximo al adiabático y que las temperaturas centrales se encuentran en el dominio de los 20.000 a 30.000 grados Kelvin. A tales temperaturas, una mezcla de hidrógeno y helio no se solidifica. El hidrógeno metálico dentro de Júpiter y Saturno es, pues, líquido. A un nivel intermedio, donde la presión alcanza los tres millones de atmósferas terrestres, el líquido metálico pasa bruscamente a líquido molecular. A niveles todavía más altos, el líquido molecular se transforma gradualmente en gas molecular: la atmósfera de Júpiter y Saturno.

Los modelos del enfriamiento de Júpiter a lo largo de la historia del sistema solar dan cuenta de toda la potencia que el planeta está irradiando hoy por encima de la que absorbe. Modelos semejantes, aplicados a Saturno, son, empero, incapaces de explicar aproximadamente un tercio de la potencia que irradia. Al parecer, Saturno tiene una fuente de energía interna no incluida en los cálculos y que está ausente en Júpiter. Hacia la mitad del decenio de 1970, Salpeter y Stevenson propusieron una explicación confirmada ahora por la información recabada por los Voyager. Un planeta que incorpora una mezcla de hidrógeno y helio tiene dos clases de energía: calorífica y gravitatoria. Si la proporción de mezcla (esto es, la relación de hidrógeno a helio) es constante en toda la mezcla, las dos clases de energía se liberan juntas en proporciones constantes. Cuando el planeta se enfría, se contrae. Pero si cambia la proporción de mezcla (por ejemplo, si el helio cae a través del hidrógeno), se libera una cantidad adicional de energía gravitatoria. Cabe suponer ese proceso en Saturno, puesto que el planeta se ha enfriado ya hasta el punto de que el helio está precipitando en la cima de la zona de hidrógeno metálico.

Salpeter y Stevenson asimilan el proceso al de la lluvia ordinaria. Cuando se enfría una partícula de la atmósfera terrestre por debajo de su punto de saturación, el agua se condensa y las gotas de agua caen. La condensación libera el calor que se había añadido al agua para convertirla en vapor, y el ritmo de enfriamiento atmosférico se hace menor. En Saturno, el equivalente a las gotas de agua es el helio, y se libera energía a medida que estas gotas rozan con el hidrógeno fluido, a través del cual caen.

Este proceso comenzó en Saturno hace unos dos mil millones de años. Júpiter, que tiene mayor masa, no se ha enfriado todavía hasta el punto en que el planeta esté, en algún nivel, saturado de helio. Tal vez está precisamente ahora alcanzando ese punto.

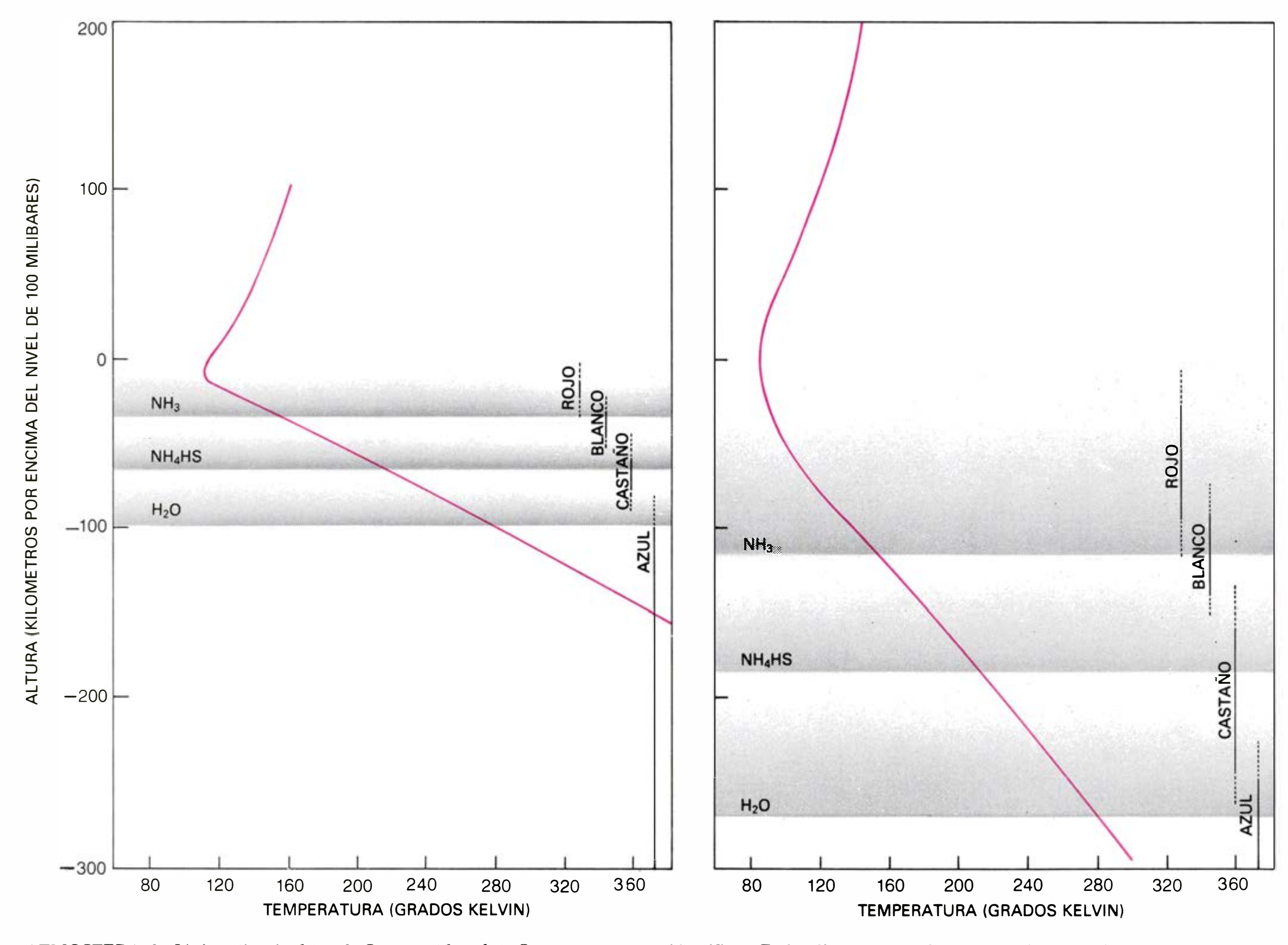
Hay pruebas que apoyan la hipótesis de que una lluvia de helio libera energía en Saturno: las que aportan las medidas de la abundancia atmosférica de helio. No es fácil obtener esos datos. El helio no absorbe radiación en el infrarrojo, lo que significa que el espectro de radiación emitido por Júpiter y Saturno no revela directamente la presencia de helio. Esta, sin embargo, afecta a la absorción de radiación infrarroja por el hidrógeno. Por este procedimiento, un grupo dirigido por Rudolph A. Hanel, del Centro de Vuelos Espaciales Goddard de la NASA, y Daniel Gautier, de la Universidad de París, ha determinado los números relativos de átomos de helio y moléculas de hidrógeno a partir de medidas realizadas por los Voyager.

Por encima de la zona de hidrógeno metálico, Júpiter resulta tener un 10 por ciento de helio, valor poco diferente del 11 por ciento de helio en el Sol. Saturno no alcanza el 6 por ciento de helio. Si el helio que falta ha precipitado hacia abajo, a través de la zona molecular, hasta muy dentro de la zona metálica, el calor que se habría liberado podría muy bien haber mantenido la fuente de calor interno de Saturno a su valor actual durante los últimos mil millones de años. En consecuencia, las deducciones sobre la composición en volumen, la estructura interna y la historia del enfriamiento de los planetas gigantes parecen concordar entre sí.

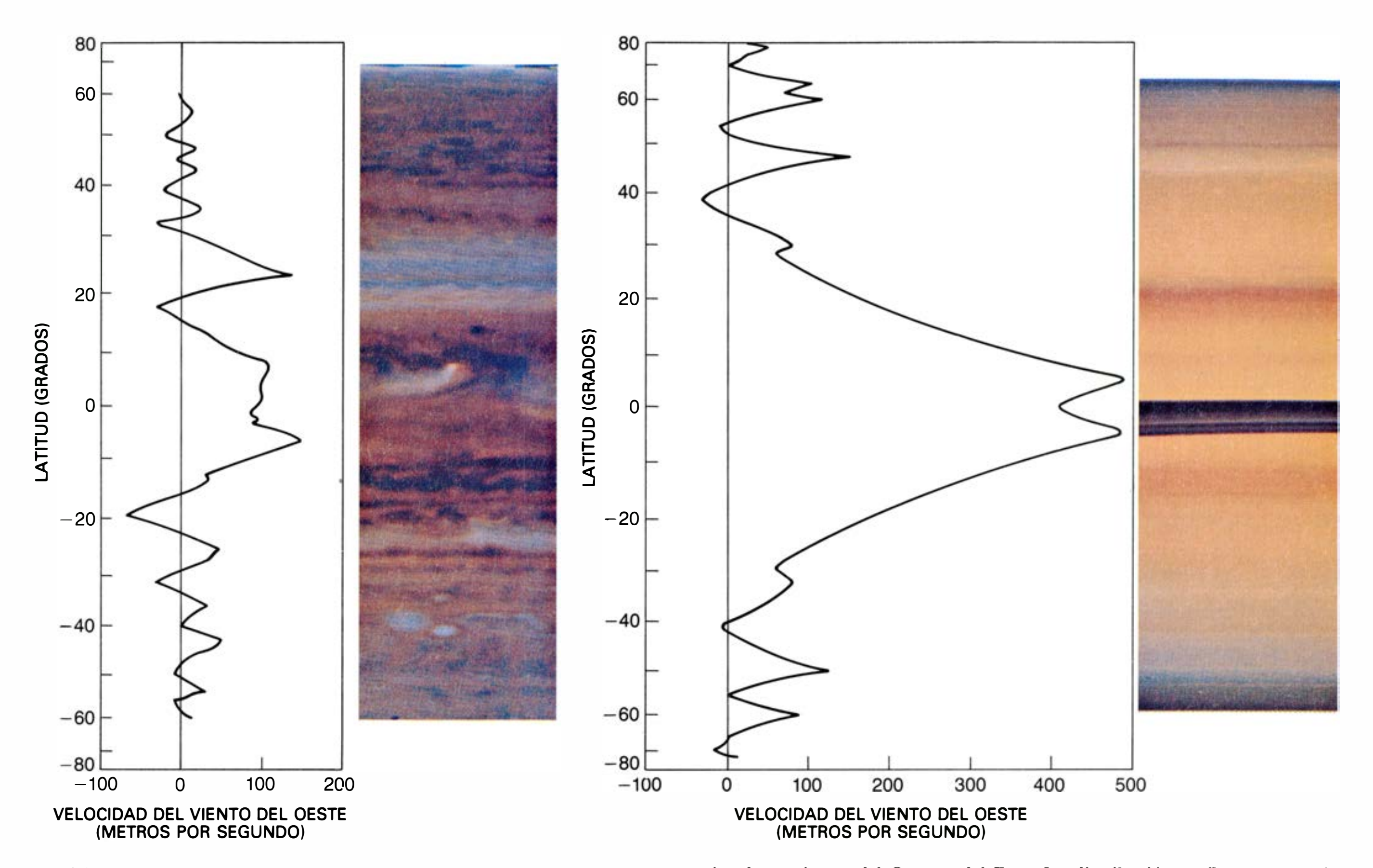
Química atmosférica

Por la década de los 30 había ya investigadores que trabajaban en la identificación de líneas en el espectro de la luz solar reflejada por Júpiter debidas a la absorción de la luz por metano y amoniaco gaseosos. La molécula de hidrógeno sólo tiene débiles líneas de absorción; de ahí que su presencia tardara 30 años en confirmarse, aunque su abundancia era, según se dedujo entonces, 1000 veces mayor que la del metano y amoniaco. Las abundancias calculadas de hidrógeno, carbono y nitrógeno en las atmósferas de Júpiter y Saturno apoyan la hipótesis de que los planetas gigantes están constituidos por materia muy parecida a la que se condensó para formar el Sol.

Sin embargo, en una mezcla de composición solar se deberían también detectar otros gases; y, por destacar los más notables: el vapor de agua y el sulfuro de hidrógeno (H_2S). En el decenio de 1970, Harold P. Larson y Uwe Fink, de la Universidad de Arizona, detectaron líneas espectrales debidas a la absorción de radiación infrarroja por vapor de agua en la atmósfera de Júpiter.



ATMOSFERA de Júpiter (*izquierda*) y de Saturno (*derecha*). Se supone que contienen proporciones de los diversos elementos químicos muy similares a las proporciones existentes en el Sol; sobre esa base, puede suponerse que cada atmósfera incluye nubes estratiformes de amoniaco (NH_3), hidrosulfuro amónico (NH_4HS) y agua. Todos esos compuestos forman partículas blancas; de aquí que los colores en cada atmósfera deben tener otras causas, todavía por identificar. En los diagramas, cada color se asigna a un intervalo de alturas de acuerdo con medidas de las temperaturas de las cimas de las nubes. Las rojas, por ejemplo, son las más frías y, por ende, las más altas. Las cimas de las nubes están a una altura no mayor que el nivel al que la temperatura (*curva en color*) empieza a aumentar con la altura, de suerte que las corrientes de convección no pueden ya transportar partículas sólidas (o líquidas) hacia arriba.



SE COMPARAN las velocidades medias del viento en Júpiter (*izquierda*) y Saturno (*derecha*) con imágenes de cada planeta. Los números positivos corresponden a velocidades dirigidas hacia el Este; los negativos, a velocidades hacia el Oeste. Las cifras son velocidades del viento respecto a la velocidad de rotación; se midieron siguiendo el movimiento de características determinadas de las nubes en imágenes sucesivas del planeta hechas por el *Voyager 2*. Resulta de las medidas que cada hemisferio de Júpiter tiene varias alternancias de corrientes del Oeste y del Este. La distribución en Saturno es más sencilla, pero más fuertes las corrientes; en la atmósfera de éste, 500 metros por segundo representan dos tercios de la velocidad del sonido. Ochenta años de observaciones y las recientes mediciones desde los vehículos espaciales sugieren que las coloraciones de las nubes a cualquier latitud pueden cambiar, si bien los vientos son más persistentes. La banda oscura a través del ecuador de Saturno es el sistema de anillos del planeta y su sombra sobre la superficie.

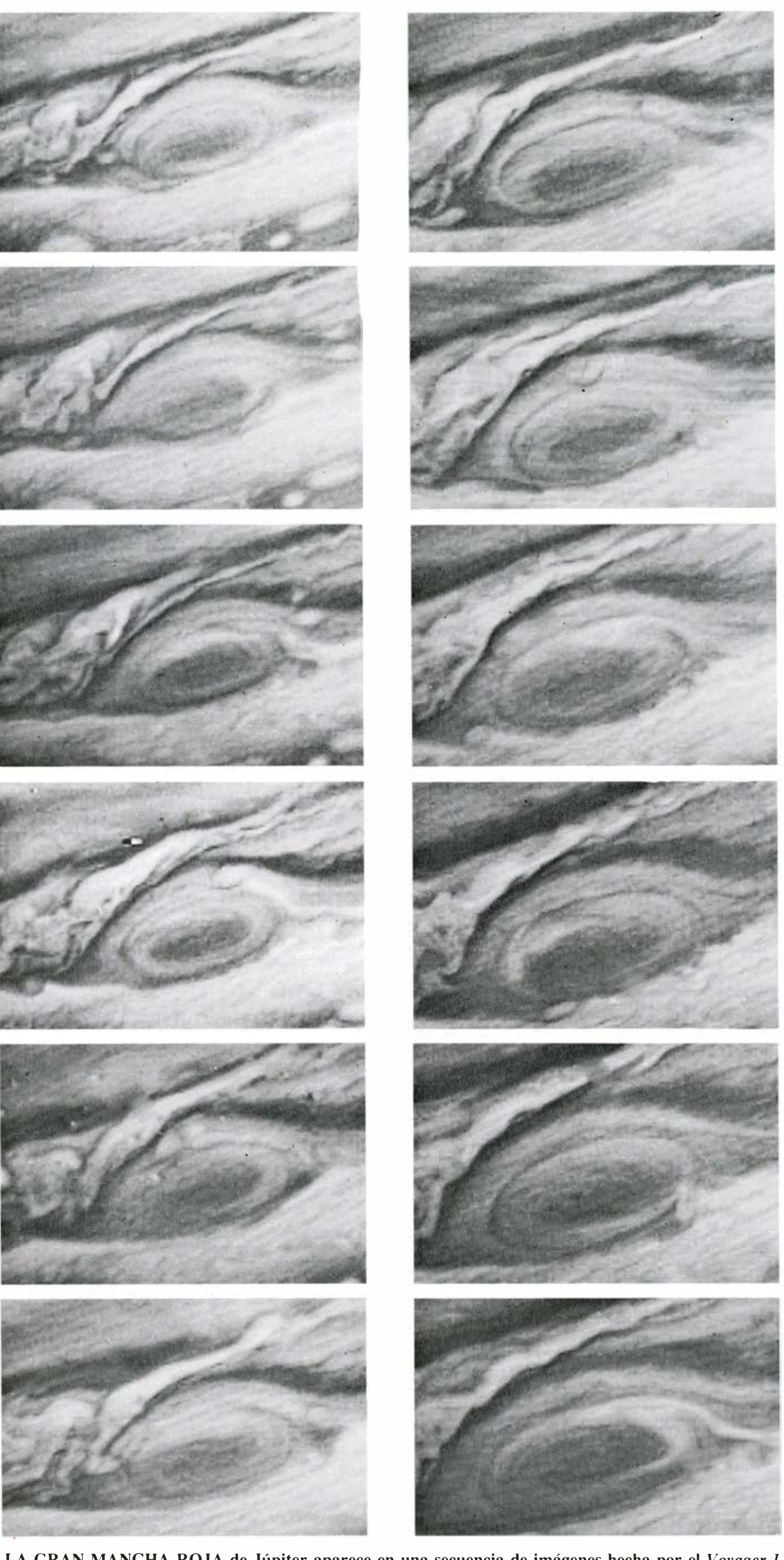
Lo que no ha ocurrido a propósito de Saturno. Ni en este planeta ni en el anterior se ha detectado sulfuro de hidrógeno. Pese a ello, tanto el vapor de agua como el sulfuro de hidrógeno gaseoso deben tener una concentración bastante baja en la cima de las nubes de Júpiter y Saturno, donde la temperatura es de 150 grados Kelvin (−123 grados Celsius) o inferior. A fin de cuentas, nada impide la presencia de esos gases, en abundancias solares, por debajo de las nubes.

Los espectros infrarrojos de las manchas calientes de Júpiter (agujeros en las nubes) revelan una rica mezcla química. Además de vapor de agua, se ha detectado fosfino (PH_3), germano (GeH_4), cianuro de hidrógeno (HCN) y monóxido de carbono (CO), así como metano pesado, esto es, moléculas de metano que contienen el isótopo pesado del hidrógeno (hidrógeno 2, llamado deuterio) o bien del carbono (carbono 13). Se ha descubierto también la existencia de etano (C_2H_6) y acetileno (C_2H_2) en todo el disco de Júpiter y en la superficie de Saturno. La mayoría de estos compuestos faltarían en una atmósfera rica en hidrógeno de hallarse ésta en equilibrio químico. Los diversos compuestos de carbono volverían a convertirse entonces en metano, y el nitrógeno formaría amoniaco.

La fuente obvia del desequilibrio químico es el Sol. En particular, la radiación ultravioleta del astro puede romper las especies químicas dominantes, así el metano, y liberar radicales libres, como CH_2 y CH_3. Idéntico efecto producen los relámpagos en las nubes de Júpiter y Saturno y el impacto de las partículas cargadas eléctricamente. (Las partículas cargadas "llueven" sobre la atmósfera procedentes de la magnetosfera, o región situada encima del planeta donde están confinadas por el campo magnético del mismo.) Los radicales pueden reaccionar con moléculas de metano para formar etano y acetileno y liberar hidrógeno. Por tanto, la composición de la atmósfera de Júpiter refleja un equilibrio entre la producción y la destrucción de esos hidrocarburos superiores.

La presencia del monóxido de carbono en las manchas calientes de Júpiter constituye un problema más arduo, ya que el agua, fuente potencial del oxígeno, escasea por encima de las nubes, donde abundan fotones ultravioleta (cuantos de radiación ultravioleta). Cabe imaginar, por otro lado, agua abundante por debajo de las nubes, y combinarse allí con el metano para formar monóxido de carbono. Las condiciones de presión y temperatura que reinan bastante por debajo de las nubes favorecen tal reacción. Otra posibilidad es que el oxígeno situado sobre las cimas de las nubes no provenga de Júpiter, sino del dióxido de azufre eyectado por los volcanes de su satélite Io, cuya actividad se reveló en las imágenes tomadas por el *Voyager 1*.

Las partículas sólidas y líquidas que constituyen las nubes de Júpiter y Saturno suministran ulterior evidencia del desequilibrio químico. Los vapores condensables más abundantes en una mezcla de composición solar son los de agua, amoniaco y sulfuro de hidrógeno. Bajo condiciones de equilibrio químico, forman cristales de hielo, amoniaco e hidrosulfuro amónico (NH_4HS). Pueden concebirse también gotas líquidas de agua y soluciones de amoniaco en



LA GRAN MANCHA ROJA de Júpiter aparece en una secuencia de imágenes hecha por el *Voyager 1* cada cuatro revoluciones del planeta, es decir, una vez cada 40 horas. La mancha misma se ha observado desde la Tierra desde hace tres siglos. Gira en sentido antihorario entre una corriente del Este al Norte de ella y otra del Oeste al Sur de la mancha. Mide unos 25.000 kilómetros de longitud. En la secuencia del Voyager se observan detalles de estructuras de hasta 1000 kilómetros. Hay pequeñas nubes acercándosele desde el Este. La rodean en 6 a 10 días. Parte de ellas son engullidas por la mancha. Las nubes más pequeñas regresan al Este. Además de la Gran Mancha Roja, Júpiter tiene varios óvalos blancos.

agua, como John S. Lewis y sus colegas, del Instituto de Tecnología de Massachusetts, han hecho observar. La dificultad estriba en que todos esos productos de condensación son blancos, mientras que las nubes de Júpiter y Saturno son coloreadas.

Ronald G. Prinn, del MIT, ha sugerido algunos posibles agentes colorantes. Por un lado, el azufre molecular (S_n, donde n puede tomar varios valores mayores que la unidad) forma partículas de color pardo y amarillo. Resulta, pues, todavía más desolador el hecho de que el sulfuro de hidrógeno, molécula padre del azufre elemental, no se haya detectado. El problema con que se enfrentan los teóricos que apoyan la idea del azufre como agente colorante es de qué manera ocultar al análisis espectroscópico el sulfuro de hidrógeno en las nubes y dejando expuesto a los fotones solares el suficiente para que se pueda formar S_n. Prinn observa que el fósforo elemental tiene un color rojizo semejante al de la Gran Mancha Roja; podría originarse en la colisión entre fotones ultravioleta y moléculas de gas fosfino. Dada la mutua afinidad de los átomos de carbono, los compuestos orgánicos complejos constituyen también una posible fuente de color. Desgraciadamente, la identificación espectroscópica de los compuestos en estado sólido no es tarea fácil, porque el conjunto de vibraciones características (y la absorción de ciertas longitudes de onda de la luz asociada a ellas) queda enmascarado por los choques entre las moléculas vecinas. De aquí que las fuentes de color en Júpiter y Saturno sigan siendo inciertas.

Pero sí podemos afirmar que las nubes de distinto color, en uno y otro planeta, están asociadas a distintos niveles de su atmósfera respectiva. Esta correlación se pone de manifiesto al comparar imágenes obtenidas en luz visible con las sacadas en radiación infrarroja. Las imágenes visibles muestran los colores; las infrarrojas distinguen las nubes frías (y por tanto altas) de las calientes (y por tanto bajas). En Júpiter, las cimas de las nubes más altas son rojas; las inmediatamente inferiores son blancas, las que siguen en altura son pardas, y las más bajas de todas (o tal vez la propia atmósfera subyacente a las nubes) son azules. Presumiblemente, los diversos compuestos responsables de esos colores se forman a diferentes niveles en respuesta a distintas temperaturas y cantidades de luz solar. Las partículas blancas que forman los componentes más importantes de las nubes deben estar también distribuidas

en capas. El amoniaco se situará en la parte más alta, después el hidrosulfuro amónico y finalmente el agua.

Circulación atmosférica

En la atmósfera de la Tierra, los gradientes horizontales de temperatura constituyen el depósito de donde los vientos obtienen su energía. Los gradientes se producen porque el Sol calienta los trópicos más que los polos; entonces el aire tropical caliente asciende sobre el aire polar más frío. Esto convierte energía potencial gravitatoria en energía cinética y también transporta calor hacia arriba y hacia los polos.

En Júpiter y Saturno los gradientes horizontales de temperatura pueden ser menos importantes que en la Tierra. En primer lugar, el gradiente de temperatura del ecuador al polo es pequeño en Júpiter y Saturno, al menos en los niveles de la cima de las nubes donde los vehículos espaciales Pioneer y Voyager pudieron medirlo. En Júpiter, por ejemplo, la diferencia de temperatura entre el ecuador y el polo no llega a los tres grados Celsius. En la Tierra, la diferencia es 10 veces mayor y está comprimida en una distancia del ecuador al polo que es 10 veces menor.

En segundo lugar, la atmósfera de Júpiter y Saturno obtiene la mitad o más de su calor del interior del planeta. En ese aspecto, ambos se parecen a las estrellas más que a la Tierra. Según las teorías de la convección empleadas en los modelos estelares, una fuente interna de calor debe originar un gradiente adiabático de temperaturas que es el mismo según cualquier dirección radial. Además, como la parte de la atmósfera de Júpiter y Saturno que recibe luz solar tiene solamente una millonésima de la masa (y, por tanto, una millonésima de la capacidad calorífica) del interior del planeta, es de esperar que la atmósfera esté, en esencia, cortocircuitada por el interior. Viene a ser como si todos los puntos de la superficie del planeta estuvieran conectados con el centro a través de conductores muy gruesos, en tanto que otros muy finos trabaran los puntos de la superficie entre sí. De este razonamiento se deduce que los gradientes horizontales de temperatura deben ser pequeños.

Pese a ello, sobre la superficie de Júpier y Saturno cae más luz solar que en los polos. Las mediciones registradas por los vehículos espaciales Pioneer y Voyager muestran que el ritmo de emisión de radiación infrarroja es aproximadamente independiente de la latitud y mayor que el de absorción de luz solar a todas las latitudes. Por tanto, se transporta calor hacia el polo. El argumento de que el interior de cada planeta constituye un conductor grueso implica que la totalidad del fluido interior del mismo toma parte en el transporte hacia el polo. Un modelo de ordenador, que hemos ideado Carolyn Porco y yo mismo en el Cal Tech, sugiere la manera como se mantiene el transporte de calor. En respuesta a la desigual distribución de energía recibida del Sol en la superficie del planeta, se crean en el interior pequeñas diferencias de temperatura. Diferencias que modulan el ritmo al cual el calor interno llega desde abajo. Según el modelo, el decrecimiento hacia el polo del calentamiento solar de la delgada capa iluminada queda compensado por un aumento hacia el polo en la cantidad de calor interno que aflora. El transporte de calor hacia el polo en la delgada capa iluminada carece de relevancia.

Gareth P. Williams, del Laboratorio de Dinámica Geofísica de Fluidos de la Administración Nacional Oceánica y Atmosférica en la Universidad de Princeton, ha publicado un modelo basado en hipótesis enteramente diferentes. Williams comenzó con un formalismo matemático empleado en la predicción del tiempo en la Tierra. Las ecuaciones describen los movimientos que se producen en una atmósfera fluida mucho más delgada que el radio del planeta, de acuerdo con leyes físicas tales como la conservación de la cantidad de movimiento y de la energía. Williams modificó el radio y la velocidad de rotación del planeta hipotético para que caracterizasen a Júpiter o Saturno; redujo también lo que se llaman variables de disipación, que representan la velocidad con que una atmósfera pierde calor y cantidad de movimiento en presencia de determinadas fuerzas, la de rozamiento por ejemplo. La reducción de estas variables resulta viable si tenemos en cuenta que Júpiter y Saturno carecen de superficies sólidas donde pueda rozar el viento. Además, una atmósfera profunda y espesa pierde calor con extrema lentitud, como ya observaron hace 15 años Peter J. Gierasch, ahora en Cornell, y Richard Goody, de la Universidad de Harvard.

Williams supuso también que la atmósfera subyacente al nivel de penetración de la luz solar en Júpiter y Saturno no incidía en la circulación suprayacente. En otras palabras, se despreció el afloramiento de calor desde el interior o se le consideró independiente de la latitud. Se supuso, además, que la frontera inferior de la capa iluminada actuaba como una superficie sólida, no sujeta a deformación ni permeabilidad. Esta hipótesis estaría justificada si la densidad de la capa iluminada fuese substancialmente menor que la de la capa situada por debajo. Tal decrecimiento se encuentra entre el agua caliente en la superficie de los océanos terrestres y el agua más fría por debajo de aquella. Resulta, empero, menos probable encontrar esa mengua en un fluido que está transportando calor por convección desde abajo hacia arriba.

En resumen, el modelo que Williams presenta de la atmósfera de Júpiter o Saturno recuerda, en muchos aspectos, la atmósfera terrestre. Lo que no ha sido óbice para que se demostrara capaz de generar la característica más notable de la meteorología de Júpiter o Saturno: una distribución alternante de vientos del Este y del Oeste. De hecho, el modelo impone una predicción sobre la manera como se mantienen los vientos. En la atmósfera terrestre, los vientos persistentes en latitudes bajas soplan hacia el Oeste, y los vientos persistentes en latitudes medias soplan hacia el Este. Ambos se mantienen gracias a los torbellinos en gran escala que transportan cantidad de movimiento dirigida hacia el Este desde las latitudes ecuatoriales hacia otras más altas. Los torbellinos incluyen ciclones y anticiclones, responsables en parte destacada del tiempo que reina en la Tierra. En Júpiter y Saturno, las bandas de vientos de dirección opuesta son más numerosas. A pesar de ello, en el modelo de Williams compete a los remolinos inyectar energía en los vientos de Este hacia Oeste. Además, la energía que alimenta a los remolinos, sea en Júpiter y Saturno como en la Tierra, proviene en último término (según propone Williams) de la misma fuente: el gradiente de temperatura, desde el Ecuador hasta los polos, mantenido por el calentamiento solar.

Mediciones de los Voyager

Las mejores medidas de vientos en Júpiter y Saturno se han obtenido siguiendo la posición de rasgos característicos de nubes en imágenes sucesivas tomadas por los Voyager. El grupo de investigadores responsables de la realización de esas mediciones lo formábamos Reta Beebe, Garry E. Hunt, Jim L. Mitchell y el que firma este artículo. Nuestro primer objetivo, por lo concerniente a Júpiter, era definir la velocidad media del viento en cada grado de latitud, promediando las velocidades de todas las características nubosas que





MANCHA DE VIDA LARGA EN SATURNO, que es un óvalo de color castaño en el hemisferio Norte fotografiado el 23 de agosto (*izquierda*), y una rotación del planeta (10 horas), a la derecha. Tiene una cuarta parte del tamaño de la Gran Mancha Roja de Júpiter. Su rotación procede en sentido horario.

podíamos identificar a dicha latitud. En la mayoría de las latitudes, se podían seguir de 50 a 100 características distintivas de alto contraste. Saturno no tiene tantos rasgos de esos y sacamos, por tanto, menos medidas.

¿A qué resultados llegamos? Destacaré, en primer lugar, la concordancia entre las medidas basadas en las imágenes de Júpiter del Voyager y las fundadas en unos 80 años de observaciones desde la Tierra. Los cambios ocurridos a lo largo de este dilatado intervalo se diría que se produjeron principalmente en la coloración de las nubes a latitudes dadas. A veces, los vientos de Este a Oeste no eran observables desde la Tierra, pero no parece que cambiaran.

Un segundo resultado tiene que ver con la manera como los vientos alteran su velocidad con respecto a la media en una latitud determinada. En Júpiter, las velocidades medias hacia el Este no superan los 130 metros por segundo. Las desviaciones respecto a esta media, que corresponden a torbellinos, son típicamente de 20 metros por segundo. Las desviaciones presentan una tendencia sistemática. Por ejemplo, en lugares donde la velocidad media hacia el Este aumenta con la latitud, los vectores que representan la velocidad (y por tanto la cantidad de movimiento) de las desviaciones tienden a estar inclinados hacia el Nordeste. De aquí que inyecten velocidad y cantidad de movimiento hacia el Norte en latitudes donde la velocidad y cantidad de movimiento son ya mayores. Este es precisamente el mecanismo por el que los torbellinos mantienen los vientos medios en la Tierra, y es también el mecanismo que subyace en el modelo de Williams. Los aparentemente caóticos remolinos de Júpiter guardan en realidad un orden más estricto del que podría sospecharse.

Hay, sin embargo, algunas diferencias cuantitativas importantes entre Júpiter y la Tierra. Por un lado, las velocidades medias de los vientos y las velocidades medias de los remolinos son mayores en Júpiter. La energía pasa de los remolinos a los vientos permanentes de Este a Oeste a una velocidad, por unidad de superficie, 10 veces mayor en Júpiter que en la Tierra. Por otro lado, el ritmo al cual la energía calorífica queda a disposición de la atmósfera en Júpiter, para posible conversión en energía cinética de los vientos, es 20 veces menor que en la Tierra, también por unidad de superficie, porque Júpiter está más frío. Así pues, el rendimiento con el que la atmósfera convierte energía calorífica en energía cinética parece ser mucho mayor en Júpiter que en la Tierra.

A pesar de eso, hay que andar con cuidado a la hora de deducir conclusiones generales sobre el rendimiento del ciclo de energía de Júpiter o Saturno, porque esos ciclos incorporan transformaciones de energía que todavía no se han medido. Por ejemplo, una masa de aire frío situada por encima de una masa de aire cálido representa cierta cantidad de energía potencial gravitatoria, que se libera cuando el aire frío desciende. Tales descensos se producen en las nubes de Júpiter y Saturno, estando por tanto ocultas a la inspección visual. En la Tierra, el rendimiento neto del ciclo energético de la atmósfera viene a ser del 1 por ciento: aproximadamente el uno por ciento de la potencia solar absorbida alimenta los movimientos en gran escala de la atmósfera, antes de que el rozamiento en ella la convierta de nuevo en calor que la Tierra radia, devolviéndolo al espacio. El resto se envía al espacio en forma de energía radiante sin haber sido convertido siquiera en energía cinética. Se desconoce el rendimiento neto de Júpiter y Saturno.

Cilindros coaxiales

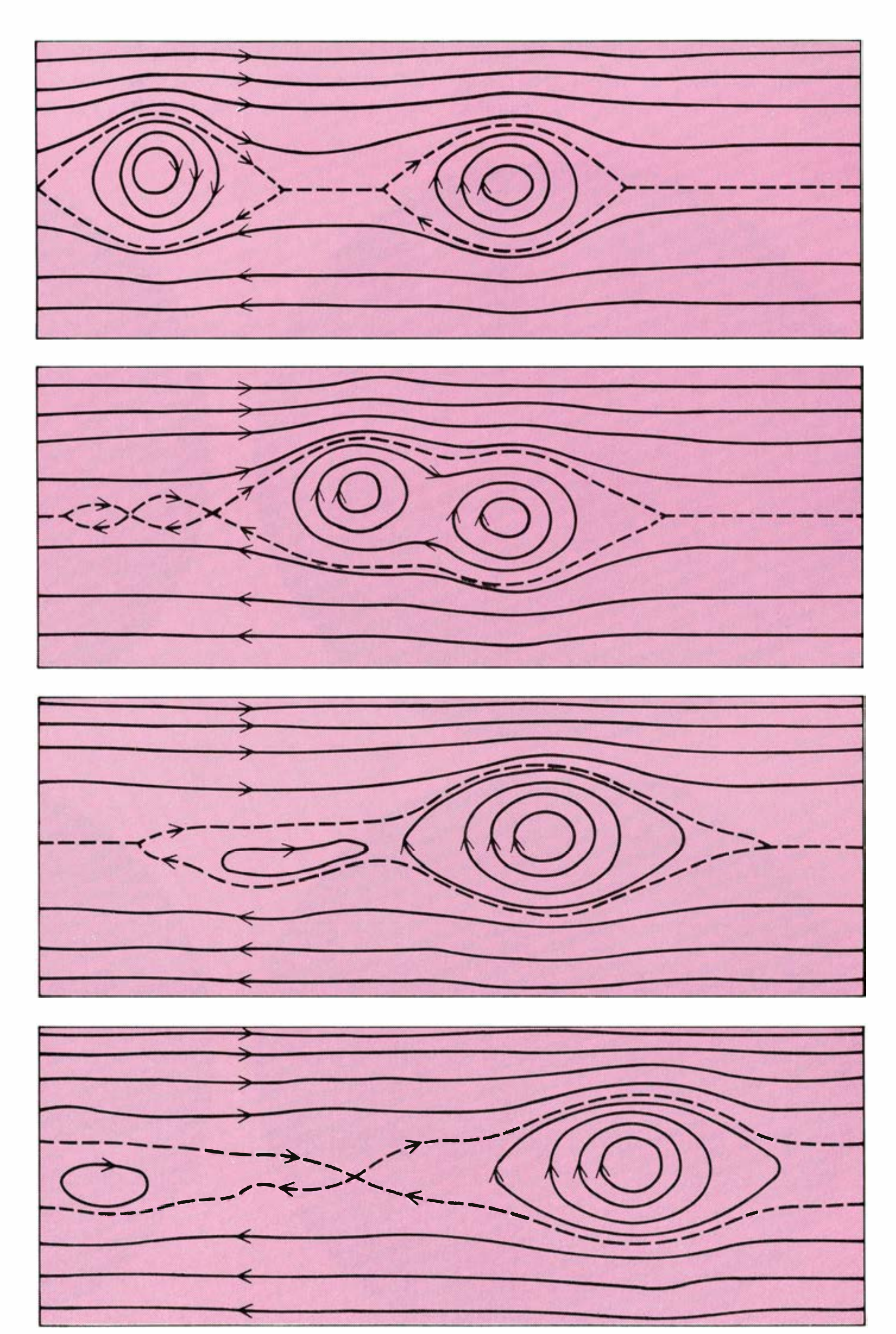
Sospecho que la gran profundidad del fluido interior no puede despreciarse a la hora de formular un modelo de la dinámica de las atmósferas de Júpiter y Saturno. Ciertos experimentos de laboratorio, incluidos los realizados hacia 1920 por Geoffrey I. Taylor en la Universidad de Cambridge y los que está acometiendo hoy F. H. Busse en la Universidad de California en Los Angeles, muestran que los movimientos turbulentos en pequeña escala (esto es, torbellinos) de un fluido en rápida rotación se alinean entre sí en columnas paralelas al eje de rotación. A cualquier nivel de cada columna, el movimiento en un plano perpendicular al eje es el mismo. Las columnas abarcan el recipiente y se resisten a ser estiradas o comprimidas a lo largo de su eje. Por

consiguiente si el recipiente es esférico, cada columna quedará a una distancia fija del eje de rotación. Los experimentos y cálculos de Busse sugieren que, como resultado de las columnas, se desarrolla en el fluido un régimen de corriente que persiste. En un líquido de densidad constante y baja viscosidad, el régimen está constituido por cilindros coaxiales de fluido que giran a distintas velocidades alrededor de su eje común, que es el eje de la rotación media del fluido. En un líquido de mayor viscosidad o con estratificación en capas de densidad variable caben otros regímenes.

La aplicación de esos experimentos de laboratorio a Júpiter y Saturno depende, pues, de la distribución de densidades en cada planeta. Un marcado aumento de densidad por debajo de la capa iluminada de la atmósfera, como postula Williams, implica que cada planeta configura la analogía a gran escala de un experimento de laboratorio con fluido estratificado. Las columnas del tipo hallado por Taylor y Busse no atraviesan la superficie de separación entre dos capas de distinta densidad; la capa iluminada puede abordarse, por tanto, separadamente de la interior. Por otro lado, la convección de calor en Júpiter y Saturno debe conducir a un gradiente adiabático de temperatura y a un interior bien mezclado. Lo que implica que cada planeta representa la analogía de un fluido de laboratorio con densidad constante.

Según la mayoría de los modelos del interior de Júpiter y Saturno, la zona adiabática acaba en las nubes. Cabe pues la posibilidad de que los vientos constantes del Este y del Oeste, que hemos medido en las nubes de ambos planetas, constituyan el signo visible del movimiento de cilindros coaxiales que se extienden a través de todo el interior fluido del planeta. Ciertamente que los vientos podrían obtener su energía de los torbellinos, como el modelo de Williams sugiere y confirman los datos de los vehículos espaciales Voyager. Pero si los vientos representan el movimiento de cilindros, la inercia que alimenta a los vientos sería inmensa. Una vez que los cilindros coaxiales se han puesto en movimiento, podrían muy bien permanecer en rotación imperturbada a lo largo de los 80 años de observaciones realizadas, primero, por astrónomos situados en la Tierra y, después, por los instrumentos introducidos en los Voyager.

Los cilindros coaxiales sí implican que los perfiles de velocidad del viento, en función de la latitud en los hemisferios Norte y Sur de Júpiter y Saturno, deben ser simétricos a uno y otro lado del Ecuador. No es preciso, sin embargo, que la simetría se extienda a altas latitudes, porque el salto de densidad en la superficie de separación entre hidrógeno molecular e hidrógeno metálico, en el interior de cada planeta, hace independientes los cilindros coaxiales internos. Para Júpiter, este fenómeno debe producirse a latitudes mayores de 40 o 45 grados Norte y Sur. Para Saturno debe afectar a latitudes superiores a unos 65 grados. En los perfiles obtenidos de los datos de los Voyager, aparecen las simetrías requeridas: los hemisferios Norte y Sur de Júpiter y Saturno muestran cada uno, aproximadamente, tres ciclos completos de vientos alternativamente del Este y del Oeste entre el ecuador y las latitudes afectadas por los cilindros desacoplados. Hay pequeñas

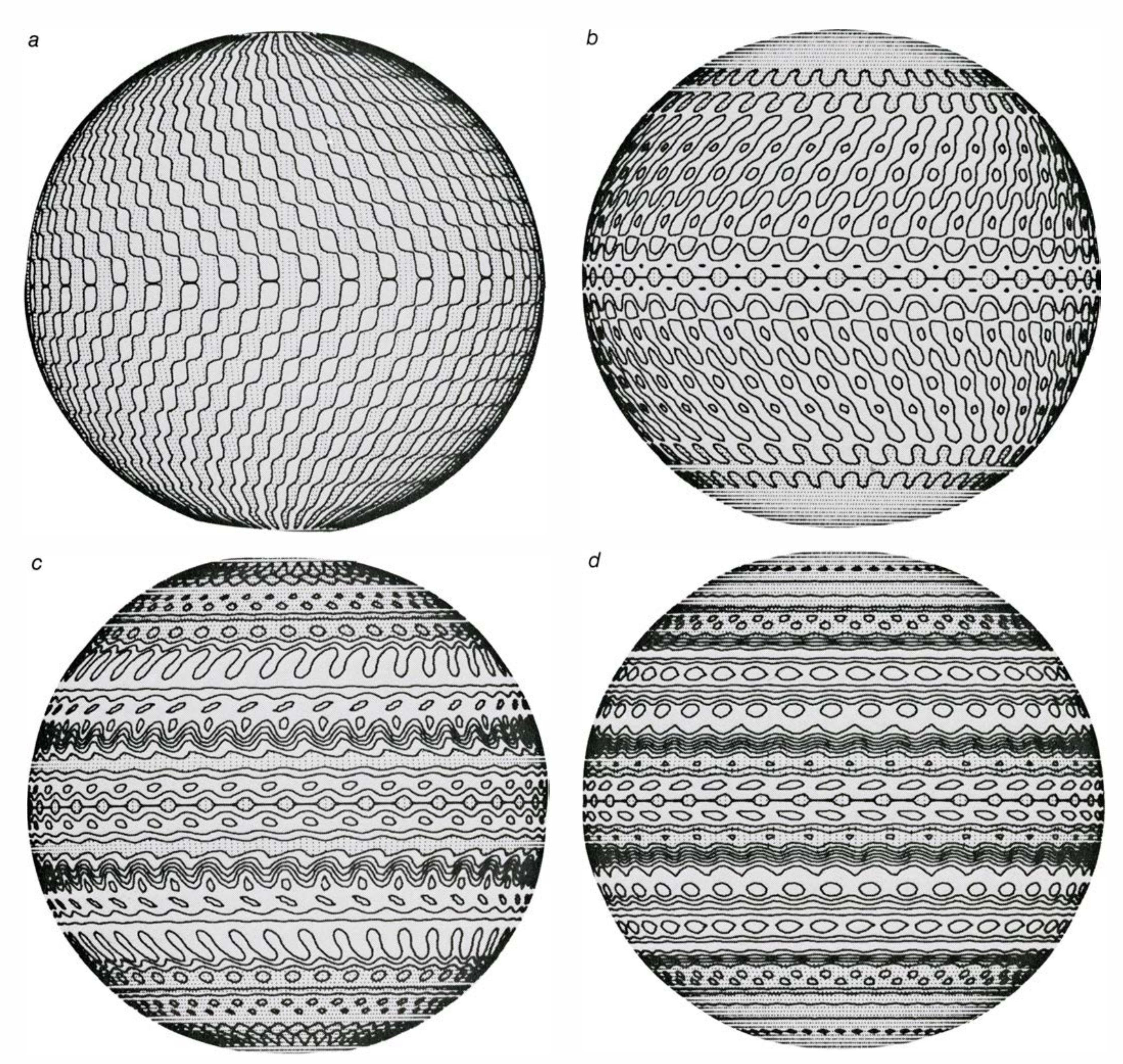


UNION DE MANCHAS en Júpiter y Saturno, simulada con un ordenador por el autor y Pham Giem Cuong, en el Instituto de Tecnología de California. Suponen que cada mancha es un vórtice más o menos permanente encima de una corriente subyacente de vientos del Oeste y del Este. Bajo esta hipótesis, las pequeñas manchas transitorias de Júpiter y Saturno poseen velocidad ascensional; esas pequeñas almacenan, pues, energía potencial gravitatoria. Las manchas grandes se mantienen engullendo a las pequeñas. Las líneas a trazos marcan los límites entre cada vórtice y la corriente laminar de la atmósfera.

desviaciones de la simetría, tales como las encontradas a 25 grados Norte y Sur en Júpiter, donde se han hallado chorros de vientos del Oeste, que quizá se deban a un gradiente de temperatura no adiabático en las capas nubosas. Desviaciones mayores implicarían un interior no adiabático, y serían prueba de que la hipótesis no se cumple.

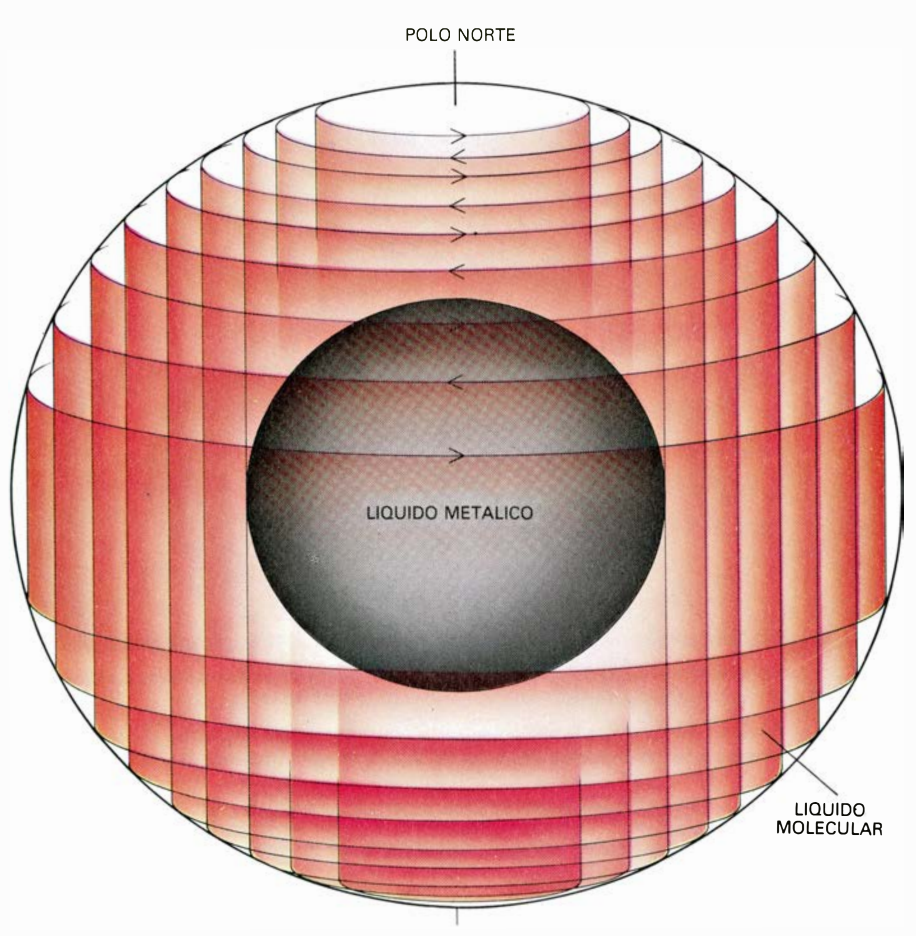
De momento, resulta difícil elegir entre un modelo de Júpiter y Saturno donde no importe la gran profundidad de fluido interior y otro donde tales profundidades sean cruciales. Después de todo, hasta los propios datos registrados por los vehículos espaciales Voyager se refieren sólo a las cimas de las nubes. El vehículo espacial *Galileo*, cuya fecha prevista de lanzamiento es a finales del decenio de 1980, explorará más profundamente la atmósfera de Júpiter. Es de esperar que la sonda pueda medir vientos al nivel de las bases de las nubes y determinar si la zona adiabática se extiende hasta esa altura. Mientras tanto, bueno será recurrir a algunos métodos indirectos.

Sea uno de ellos el que se funda en las diferencias entre Júpiter y Saturno. En particular, los vientos a la altura de las cimas de las nubes de Saturno son tres o cuatro veces más fuertes que los de Júpiter. (Las velocidades del viento del Oeste en el ecuador de Saturno son de casi 500 metros por segundo, es decir, 1800 kilómetros por hora.) Las corrientes del Oeste y del Este son más anchas que en Júpiter, y no se encuentran tantas macroestructuras ovaladas. Las causas de esas diferencias no pueden ser muchas. Quizás importe que la zona de hidrógeno metálico esté a una profundidad mucho mayor en Saturno que en Júpiter. El correspondiente aumento en el espesor de la zona de hidrógeno molecular podría dilatar la distancia entre cilindros coaxiales que girasen en sentidos opuestos y conferir-



MODELO PLANETARIO propuesto por Gareth P. Williams, del Laboratorio de Dinámica Geofísica de Fluidos de la Administración Nacional Oceánica y Atmosférica, que supone que los procesos que dan forma al tiempo en la Tierra también rigen el tiempo en Júpiter y Saturno. La desigual distribución de la radiación solar en el disco del planeta crea corrientes turbulentas en una delgada capa iluminada de la atmósfera. La cantidad de movimiento de los torbellinos alimenta después a las corrienes del Oeste y del Este. En el modelo se producen trenes de manchas, pero no óvalos aislados. La ilustración muestra el estado de la simulación por ordenador del profesor Williams después de 4,6 días (*a, arriba a la izquierda*), 23 días (*b*), 46 días (*c*) y 73,3 días (*d*).

les también mayores velocidades relativas. Comparada con la de Júpiter, la gravedad en la superficie de Saturno es menor; el flujo de calor, más pequeño y, mayores, los cambios estacionales. (Lo último es consecuencia de que el eje de rotación de Saturno esté más inclinado que el de Júpiter con respecto al plano de su órbita alrededor del Sol y, también, de que los anillos de Saturno arrojan una sombra sobre la superficie que cambia su posición con la estación.) Merced a la menor gravedad, las nubes de Saturno son más espesas; asimismo, en virtud del menor flujo de calor, los movimientos convectivos en pequeña escala de la atmósfera podrían ser más débiles. Las implicaciones de las diferencias en la superficie de cada planeta en lo que afecta a los movimientos en gran escala en la atmósfera no están, empero, aclaradas.



UN MODELO DISTINTO, avanzado por F. H. Busse, de la Universidad de California en Los Angeles, y apoyado por el autor, supone que los vientos del Este y del Oeste, visibles en las cimas nubosas de Júpiter y Saturno, constituyen la señal visible de un sistema de rotaciones que se extiende a todo el fluido interior de cada planeta. El modelo se basa en los experimentos de Geoffrey I. Taylor, de la Universidad de Cambridge, y los de Busse, que parecen indicar que en un planeta fluido en rotación, que está bien mezclado por la convección, los movimientos sostenidos son los de cilindros fluidos coaxiales. Una discontinuidad de la densidad del interior del planeta en la parte superior de la zona del hidrógeno metálico interrumpe los cilindros más interiores. Las pequeñas asimetrías entre los hemisferios Norte y Sur en los perfiles de viento que aparecen en la página 47 no resultan incompatibles con el modelo de Busse.

Ovalos de larga vida

Otra estrategia indirecta sería la del examen de las configuraciones atmosféricas de Júpiter y Saturno, distintas de los vientos del Este y del Oeste, para ver qué hipótesis acerca del interior profundo son compatibles con las observaciones. En particular, la Gran Mancha Roja de Júpiter y otros óvalos de larga vida en éste y en Saturno constituyen características únicas, y posiblemente diagnósticas, de los planetas gigantes. Los óvalos mismos se mueven de manera relativamente lenta. Por ejemplo, la Gran Mancha Roja avanza hacia el Oeste a sólo unos metros por segundo, mientras que los vientos a su alrededor alcanzan velocidades de 100 metros por segundo. Todos los óvalos de Júpiter y Saturno parecen girar como cojinetes de bolas entre corrientes adyacentes del Este y del Oeste. Cada rotación transcurre en escasos días.

Los óvalos son también resistentes: duran decenios e incluso siglos. Los torbellinos en los océanos y en la atmósfera de la Tierra son menos persistentes, a una distancia de varios órdenes de magnitud. Los remolinos del Atlántico, por ejemplo, tienden a derivar hacia el Oeste hasta que se sueldan a la corriente del Golfo, mar adentro respecto a la costa oriental de Norteamérica. Sus vidas se miden en meses y a veces en años. Los remolinos de la atmósfera terrestre son de varias clases. Los más duraderos parecen estar confinados en puntos donde concurren ciertos accidentes de la superficie: cadenas montañosas o límites entre continentes y océanos. En Júpiter y Saturno no existe pareja topografía.

Se han avanzado al menos dos propuestas que dan cuenta de muchas de las propiedades de los óvalos. Ambas hipótesis tratan de mostrar cómo un vórtice aislado puede persistir en medio de una distribución de corrientes alternativamente del Este y del Oeste. En otras palabras, tratan de demostrar que la configuración es estable, aun cuando se produzcan pequeñas perturbaciones en ella. Las propuestas no indican cómo se producen los vórtices.

Tony Maxworthy y Larry G. Redekopp, de la Universidad de California del Sur, han avanzado que un óvalo de larga vida representa una "onda solitaria", esto es, una onda automantenida con sólo una cresta, y no un tren de crestas y senos. Tal onda constituye una curiosidad en dinámica de fluidos, que se remonta hasta el siglo pasado; en la hipótesis de Maxworthy y Redekopp es un único corrimiento Norte-Sur de líneas de corriente que, por lo demás, están orientadas de Oeste a Este. Los modelos que han publicado Maxworthy y Redekopp muestran una sorprendente semejanza con la Gran Mancha Roja. Además, esos modelos implican que una onda planetaria solitaria puede existir sólo en una distribución de corriente de Este a Oeste, que es inestable respecto a ciertas perturbaciones. Esto puede explicar por qué los óvalos aislados no aparecen espontáneamente en los modelos de Williams, donde las corrientes del Este y del Oeste no eran inestables. En los modelos de Williams, las analogías más próximas a óvalos aislados son trenes de manchas dispuestos a latitud constante. Tales configuraciones aparecen en Júpiter, pero difieren de los óvalos grandes y aislados. Si dos ondas solitarias se encuentran, se cruzan sin más; pero si lo hacen dos óvalos, Júpiter y Saturno, acaban a veces por soldarse.

La segunda propuesta, que lancé con Pham Ghiem Cuong, del Instituto de Tecnología de California, parte de la hipótesis de que la distribución de corrientes del Este y del Oeste en las nu-

bes se inscribe en una distribución mucho más profunda, tal vez de cilindros coaxiales en rotación. Bajo tal hipótesis puede demostrarse que los vórtices estables existen en una corriente Este-Oeste que sea también estable. Afirma nuestra propuesta que tal vórtice se extiende hacia abajo solamente hasta la cima de la zona adiabática, dondequiera que ésta se encuentre. Sometimos a prueba la estabilidad de los óvalos en nuestro modelo de ordenador, introduciendo perturbaciones grandes y pequeñas, provocando el choque de dos óvalos e inyectando pequeños óvalos en los grandes. Dada la corriente correcta de Este a Oeste por debajo de ellos, los óvalos son muy robustos: sobreviven a perturbaciones considerablemente grandes. Además, las manchas grandes pueden crecer absorbiendo a las más pequeñas. Es probable que en Júpiter y Saturno las manchas transitorias, más pequeñas, obtengan su energía de su fuerza ascensional.

Hasta ahora, el modelo de Williams de Júpiter y Saturno, que se basa en modelos de la Tierra, es la única propuesta completa en el sentido de que posee fuentes y sumideros de energía y las corrientes del Este y del Oeste se producen espontáneamente. En grado variable, los otros modelos dan por supuesta una gran parte de la distribución. Por ejemplo, los modelos de los óvalos de larga vida presuponen las corrientes básicas del Este y del Oeste.

Pero los problemas fundamentales siguen pendientes. ¿Hasta qué profundidad se extienden las configuraciones visibles de corrientes en Júpiter y Saturno? ¿Cómo está estratificada la atmósfera, en función de la densidad, por debajo de las cimas de las nubes? ¿Cuál es la importancia relativa del calentamiento solar de la atmósfera en comparación con el calentamiento interno? Se trata de cuestiones de especial acicate, habida cuenta de la inviabilidad de un modelo de ordenador de Júpiter y Saturno que comprenda todos los factores. Simplemente, no se pueden incorporar en la misma descripción matemática de un planeta los torbellinos pequeños y transitorios, que son cruciales en la hipótesis de Williams, y las respuestas, lentas y de gran escala, al desigual calentamiento solar que se produce en las hipótesis avanzadas por Porco y por el autor. Las escalas de tamaño y tiempo de los dos tipos de fenómenos son, lisa y llanamente, muy diferentes. Los modelos de ordenador mantendrán su interés, pero a medida que vayamos comprendiendo los planetas gigantes, la intuición y la claridad de ideas se harán todavía más necesarias.



El oído de la lechuza

El ave explota las diferencias entre los sonidos que captan sus oídos izquierdo y derecho para atrapar ratones en la oscuridad. Localiza sonidos con mayor precisión que ninguna otra especie investigada

Eric I. Knudsen

La vida de la lechuza depende de su oído. Cazador nocturno, debe descubrir ratones de campo aprovechándose sólo de los crujidos y sonidos chirriantes que señalan su paso por la nieve o la hierba. Como los depredadores que cazan piezas del suelo, la lechuza debe localizar con rapidez y precisión a su presa en el plano horizontal; y, al hacerlo desde el aire, tiene que determinar, asimismo, su ángulo de elevación sobre la presunta víctima. La lechuza ha resuelto con éxito el problema: localiza sonidos en acimut (la dimensión horizontal) y en elevación (la dimensión vertical) mejor que cualquier otro animal de cuya capacidad auditiva existan datos.

¿Cómo se explica esta agudeza? La respuesta reside en la capacidad que tiene la lechuza de distinguir sutiles diferencias entre el sonido que llega a su oído izquierdo y el que le llega al derecho. Los oídos se hallan, por lo general, a distancias ligeramente distintas de la fuente sonora: las ondas acústicas se recibirán, pues, con un leve desfase temporal. La lechuza, particularmente sensible a esas diferencias minúsculas, las explota para determinar el acimut del sonido. El oído más próximo a la fuente percibe, además, una intensidad acústica mayor del mismo ruido, diferencia de la que el ave sacará partido para calibrar la localización horizontal. Y no sólo para esto: debido a una insólita asimetría en los órganos de la audición, la diferente intensidad percibida le valdrá para especificar a qué altura se halla. El oído derecho y su abertura se dirigen ligeramente hacia arriba; el oído izquierdo, y la suya, hacia abajo. Lo que explica que el oído derecho sea más sensible a los sonidos que proceden de arriba, y el izquierdo a los que se originan abajo.

El desfase en el tiempo y la diferencia en la intensidad proporcionan al ave información suficiente para localizar sonidos tanto en el sentido horizontal como en el vertical. Pero de nada le servirían esos datos si la lechuza no los organizara e interpretara. Buena parte del procesamiento se efectúa en centros cerebrales próximos al punto de arranque del canal auditivo. Desde esos centros, los impulsos nerviosos viajan hasta una red de neuronas del cerebro medio, cuya disposición semeja un mapa viario. Cada neurona de la red se excita únicamente por los sonidos procedentes de una pequeña fracción del espacio. Desde esta estructura, los impulsos se transmiten a los centros superiores del cerebro. La selección de los datos sensoriales y su transformación en un mapa espacial permiten que la lechuza localice a su presa, en oscuridad total, con una precisión implacable.

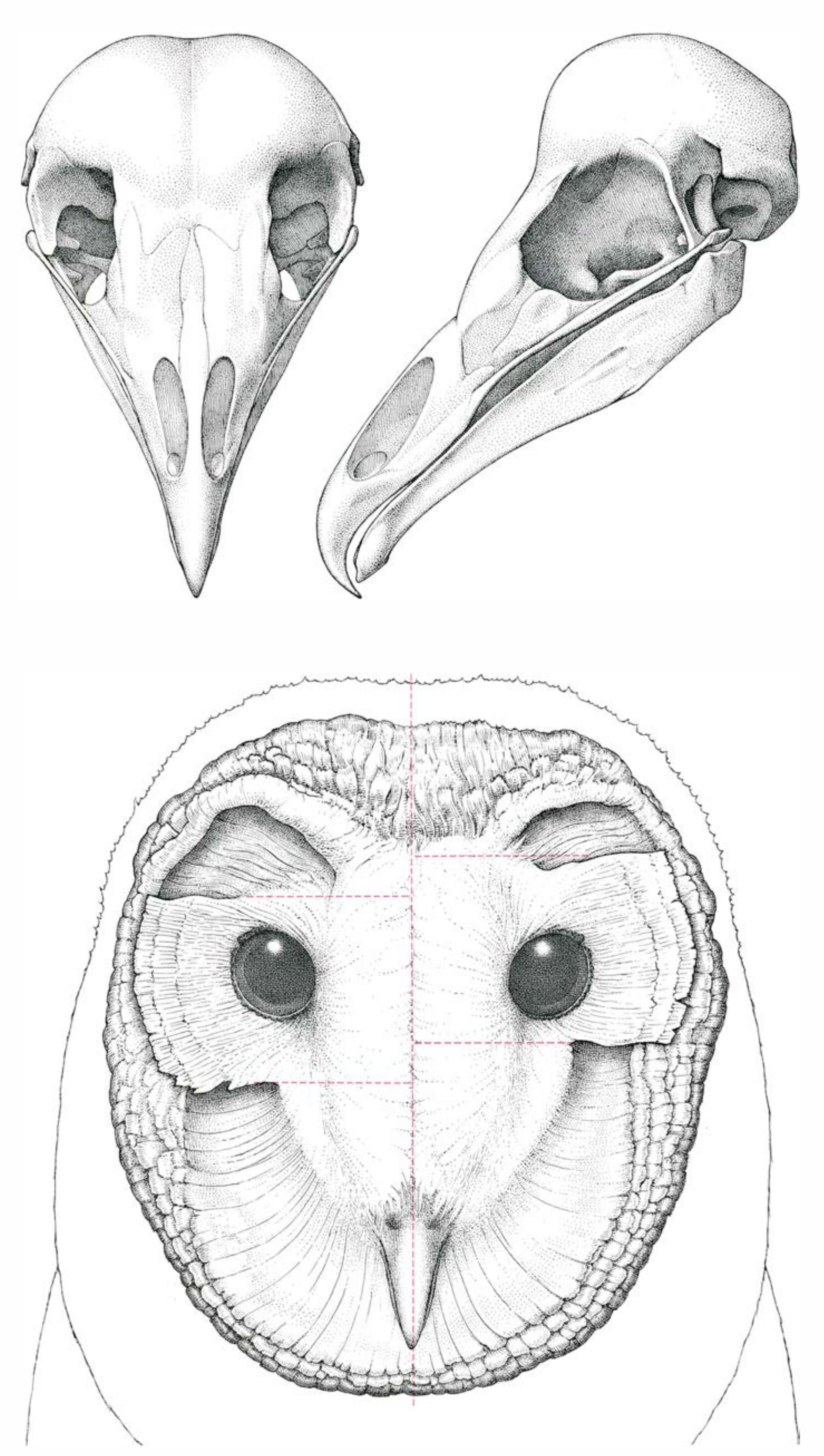
La lechuza, que goza de una amplia distribución geográfica, no circunscribe su caza a un solo tipo de presa. Hay lechuzas en todas las zonas tropicales y templadas del mundo. Se las ve cerca de las casas, anidando en graneros y campanarios. También ponen sus crías en huecos de los árboles y en agujeros practicados en taludes de tierra o en roca. Como la mayoría de las Estrigiformes, viven en parejas durante largos períodos, a veces toda la vida, retornando al mismo lugar, para reproducirse, un año tras otro. Cazan en zonas abiertas, y cubren más terreno que ninguna otra ave nocturna. Estudios de sus egagrópilas (pequeños ovillos regurgitados por la lechuza y que contienen los restos indigeribles de las presas) han demostrado que más del 95 por ciento de sus presas son pequeños mamíferos, principalmente ratones de campo; el resto se compone de anfibios y otras aves.

Las nueve especies de lechuza poseen suficientes rasgos propios, que las distinguen de mochuelos y búhos, como para formar una familia taxonómica: Titónidos. La lechuza común, *Tyto alba*, es la especie más abundante. Mide de 30 a 45 centímetros de altura y tiene la cara blanca, el dorso anteado y el pecho anteado sobre blanco; en su parte inferior domina el blanco con lunares negros. Los dedos medios del ave están provistos de un pequeño peine con el que arregla sus plumas.

El rasgo anatómico más llamativo es la cara; también es el que desempeña el papel más importante en la localización de la presa. Ancha y redonda, la cara está constituida principalmente por capas de plumas rígidas y densas dispuestas en hileras muy apretadas. Esta estructura plumosa, denominada gola o gorguera facial, forma una superficie que refleja con gran eficacia los sonidos de alta frecuencia. El cráneo es bastante estrecho y pequeño.

A lo largo de la gorguera corren dos conductos, desde la frente hasta la mandíbula inferior; ambos miden unos dos centímetros de ancho por nueve de largo. La forma de los conductos es similar al pabellón externo carnoso del oído humano, y cumplen el mismo cometido que la oreja: captar sonidos de alta frecuencia procedentes de un gran volumen de espacio y encauzarlos hacia el interior de los canales auditivos. Los conductos se unen bajo el pico, pero se hallan separados por encima de éste por una gruesa cresta de plumas. Las propias aberturas auditivas se esconden bajo las haldas preaurales, dos colgajos de piel que surgen hacia los lados, jun-

ACOMETIDA DE LA LECHUZA en oscuridad total. Impresionante secuencia fotográfica obtenida en el laboratorio con radiación infrarroja, a la que no son sensibles los ojos de la lechuza. A diferencia de los depredadores que viven sobre el suelo, el ave debe determinar su elevación sobre la presa, así como la dirección en el plano horizontal. Localiza ratones con la pista exclusiva que le ofrece el ruido. Antes de su arremetida, alinea las garras con el eje longitudinal del cuerpo del ratón, como aparece en la última imagen. Esta acción demuestra que el ave puede inferir el movimiento de la presa a partir del sonido.



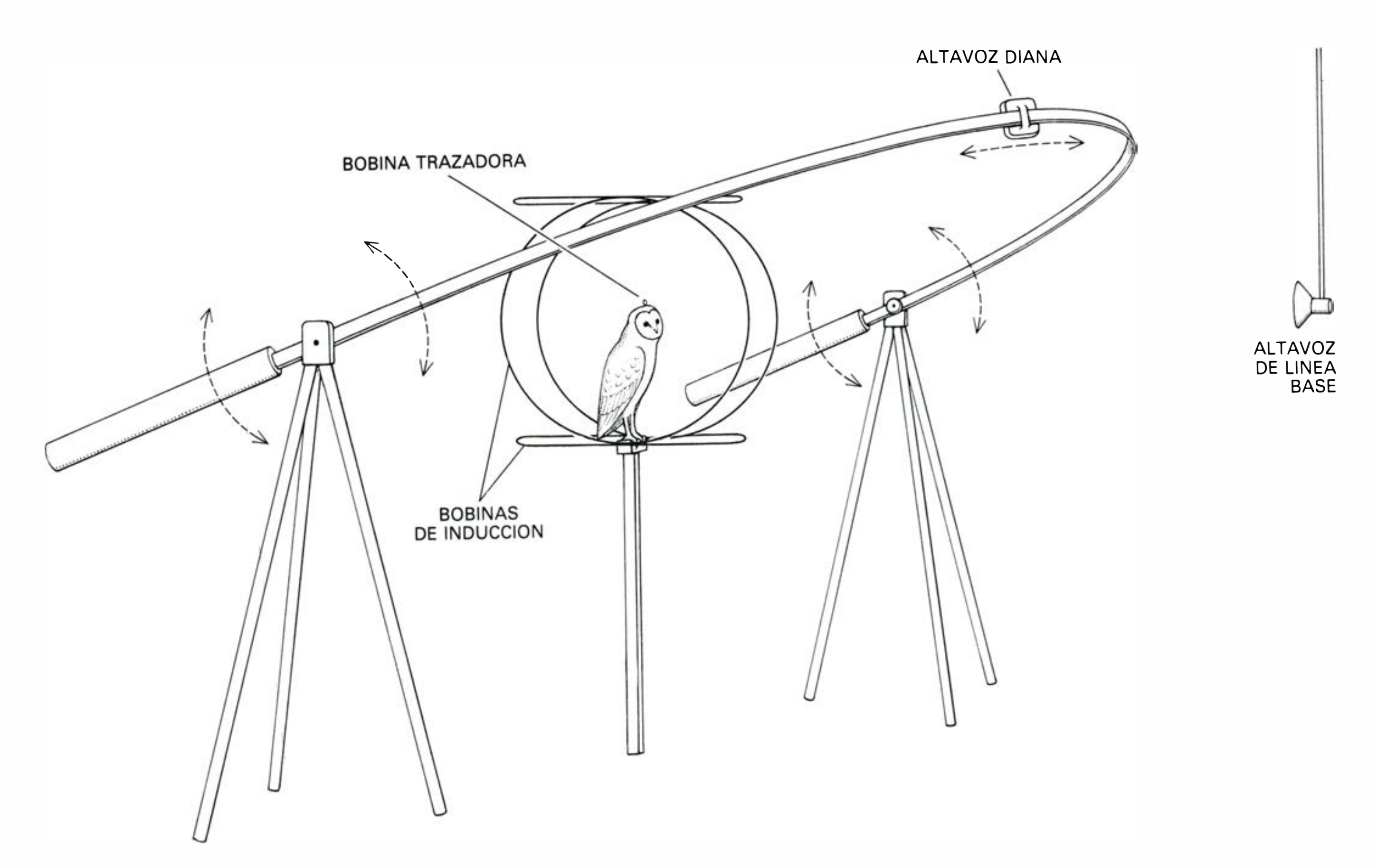
ESTRUCTURA FACIAL de la lechuza, que condiciona buena parte de su precisión auditiva. La cara está formada por filas de plumas muy apretadas, la llamada gola o gorguera facial, que se extienden desde el cráneo, relativamente estrecho. Los oídos externos son conductos formados por la gorguera que corren a lo largo de toda la cara para unirse bajo el pico. Captan los sonidos y los conducen a las aberturas del canal auditivo, ocultas bajo las haldas preaurales (colgajos de piel que salen hacia afuera desde cerca de los ojos). El oído izquierdo es más sensible a los sonidos de baja frecuencia procedentes de la izquierda, y el oído derecho lo es más a los de la derecha. Sin embargo, a altas frecuencias, el oído derecho es más sensible a los sonidos que proceden de arriba, porque la posición del halda preaural y de la abertura de ese lado es algo inferior, y el conducto está inclinado hacia arriba. La abertura y el halda se sitúan más arriba en el lado izquierdo, y el conducto está inclinado hacia abajo. El oído izquierdo es más sensible a sonidos que proceden de abajo. De esta forma, las diferencias de intensidad del sonido percibido por uno y otro órgano informan sobre la elevación de la fuente de sonido, así como de su dirección horizontal.

to a los ojos. Toda esta compleja estructura facial queda velada por una capa de plumas, particularmente finas, que son transparentes al sonido. Las propiedades acústicas de la gorguera facial dan razón del método que sigue el ave para localizar el origen del sonido.

Para sobrevivir, ha de descubrir a su presa a través del ruido que evidencie su presencia. No es fácil ver un ratón de campo, ni siquiera a plena luz del día, porque su coloración se confunde con la de su entorno; además, tienden a desplazarse a través de túneles abiertos entre la hierba o la nieve. De noche, cuando salen en busca de alimento, resultan prácticamente invisibles, por penetrantes que sean los ojos de la lechuza. Cazar desde el aire torna más difícil el empeño; hay que determinar el ángulo de elevación sobre la presa. Y atender sólo al acimut supondría abandonar toda una línea de posibles objetivos del suelo.

La lechuza localiza sonidos, con gran precisión, en dos dimensiones del espacio. Roger S. Payne, y posteriormente Masakazu Konishi y yo, demostramos que identificaba el origen de un sonido en un orden de precisión de uno a dos grados, tanto en acimut como en elevación; un grado equivale aproximadamente a la anchura del dedo meñique respecto de la longitud total del brazo. A más de uno sorprenderá saber que, hasta que se mensuró la capacidad de la lechuza, se consideró al hombre la especie dotada de mayor agudeza para adscribir el origen de un sonido. La verdad es que los humanos rayan la precisión del ave en acimut, pero ésta triplica nuestras posibilidades en elevación. Monos y gatos, especies que han adquirido un oído excelente, alcanzan sólo la cuarta parte del valor que señalara un hipotético listón que midiera la capacidad de localizar sonidos en la dimensión horizontal, y cuyo tope correspondiera a la "marca" obtenida por la lechuza. (Se trata de la única dimensión sujeta a prueba en esos mamíferos.)

La sensibilidad del oído de la lechuza queda demostrada por su capacidad para ubicar sonidos distantes y su destreza para orientar las garras antes del golpe final. Cuando se abate sobre un ratón, incluso en una habitación experimental en completa oscuridad, alinea rápidamente sus garras con el eje del cuerpo de la víctima. Fue Payne quien adelantó que este comportamiento no era casual. Cuando el ratón se da la vuelta y corre en una dirección dis-



PRUEBAS DE ORIENTACION DE LA CABEZA, que medían la precisión del oído de la lechuza. Estos experimentos se basan en una respuesta natural durante la caza: al percibir un ruido, vuelve la cabeza hasta encararse hacia el origen del sonido. En este dispositivo experimental, el ave permanecía posada sobre un pedestal. Una bobina "trazadora" montada sobre su cabeza se situaba en la intersección de un campo magnético horizontal y otro vertical inducidos por bobinas inmóviles. Cualquier movimiento de la cabeza originaba un cambio medible en la corriente de la bobina trazadora. La atención de la lechuza se dirigía primero hacia un sonido que procedía de un altavoz fijo, que marcaba la línea base, o "cero". Después, un sonido procedente de un altavoz móvil, la diana, provocaba el giro de la cabeza en un movimiento rápido. Un ordenador controlaba la localización del altavoz diana y registraba los movimientos de la cabeza. Con este dispositivo podía medirse la precisión con que la lechuza respondía a sonidos procedentes de distintos puntos.

tinta, la lechuza realinea sus garras en conformidad. Es claro que este comportamiento aumenta la probabilidad de un golpe afortunado; asimismo, implica que la lechuza, además de identificar la localización de la fuente sonora con una precisión extrema, detecta también cambios sutiles en el origen del sonido, a partir de los cuales infiere la dirección del movimiento de la presa.

Experimentos de distinto diseño han contribuido a dilucidar cómo logra realizar tan difíciles hazañas. Así, en pruebas con aves en vuelo libre se ha medido la precisión con la que la lechuza vuela hacia una fuente de sonido invisible y la golpea. Experimentos de orientación de la cabeza, en los que se mantenía al ave posada sobre un pedestal de prueba, han contribuido a medir la precisión con la que alineaba su cabeza con la emisión sonora. Experimentos en los que se sondaba con un microelectrodo el cerebro de una lechuza anestesiada, mientras se dejaban oir varios sonidos, han demostrado de qué manera organiza e interpreta el sistema nervioso central la información sensorial.

Konishi y yo hemos efectuado una serie de pruebas donde el argumento principal era la orientación de la cabeza. Este sistema experimental ofrece varias ventajas en comparación con otros realizados con aves en vuelo libre. En las pruebas de vuelo libre, los errores de vuelo pueden confundirse con errores en la localización de los sonidos. Además, no puede determinarse el ángulo que forman la fuente del sonido y la cabeza del ave en el momento en que decide atacar. Ha de agregarse, además, que las pruebas de vuelo libre son complicadas y llevarlas a cabo representa una gran inversión de tiempo. En cambio, los experimentos de orientación de la cabeza, de más fácil ejecución, permiten medir la relación entre la cabeza y la fuente del sonido.

Estos experimentos se apoyan en una respuesta natural de su comportamiento depredador. Al oír un ruido, el ave vuelve la cabeza en una rápida sacudida y la orienta directamente hacia la fuente sonora. Este movimiento sitúa los sonidos en la región de mayor agudeza auditiva. Los ojos de la lechuza son inmóviles; el giro le va a permitir, pues, ver el blanco con la máxima precisión. Konishi y el autor analizaron ese comportamiento montando sobre la cabeza del animal una liviana bobina "trazadora". Se centraron los campos magnéticos generados por otras bobinas; así, cuando el ave se hallaba posada normalmente, la bobina trazadora se encontraba en la intersección del campo horizontal con el vertical. La corriente eléctrica inducida en la bobina trazadora variaba según su orientación con respecto a estos campos. Valorando la magnitud de dos señales distinguibles podían medirse la componente horizontal y la vertical de la orientación de la cabeza.

Las pruebas se hicieron en una habitación totalmente a oscuras, revestida con materiales que eliminaban el eco. Los sonidos se emitían por un altavoz inmóvil (el altavoz que marcaba la línea base, o "cero") y uno móvil (el altavoz "diana"). La lechuza se volvía primero hasta encararse con el sonido procedente del altavoz de línea base, situado frente a su posadero; después, un sonido emitido por el altavoz diana provocaba el giro de la cabeza con el caracte-

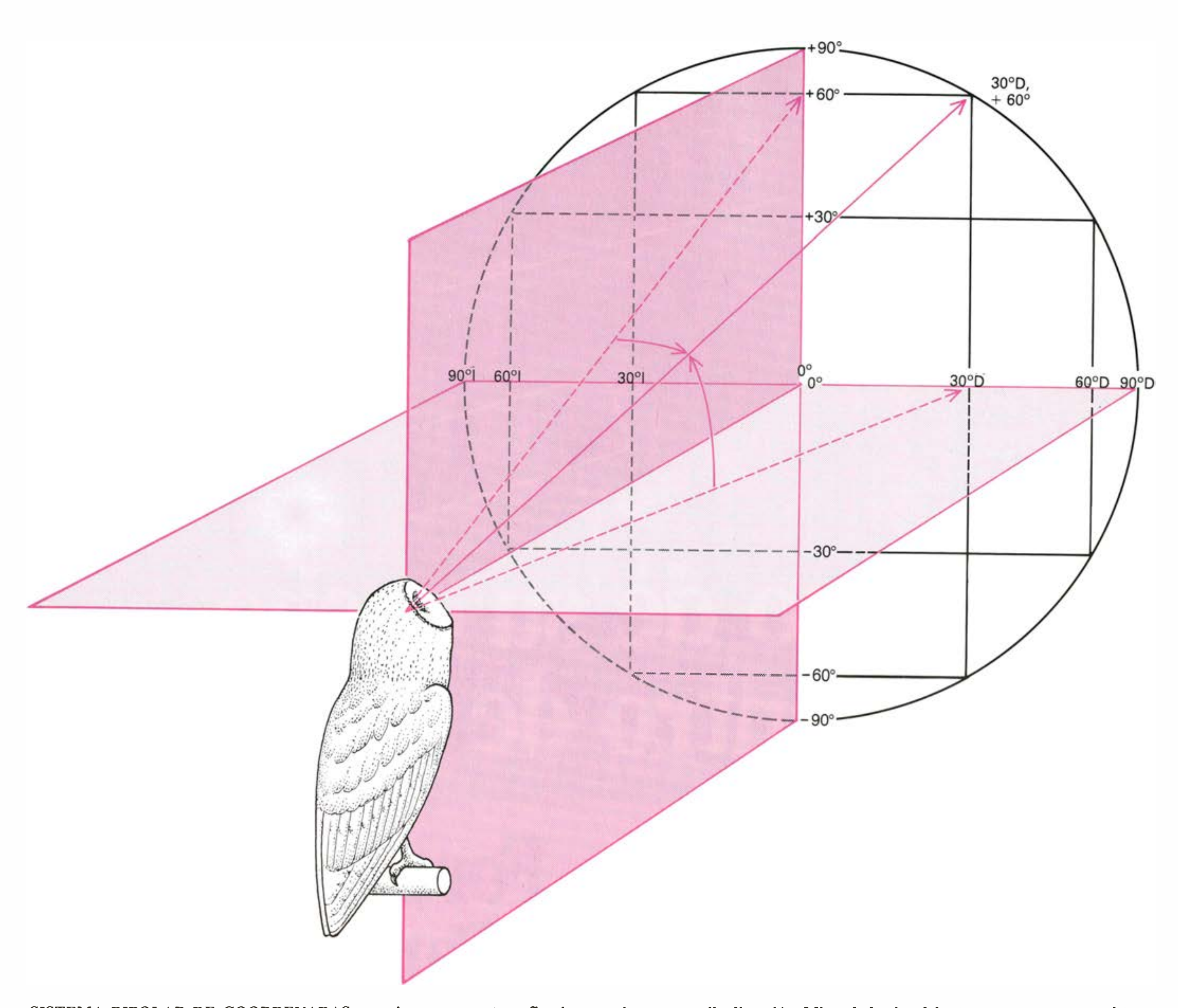
rístico movimiento de sacudida. Un ordenador controlaba la ubicación del altavoz diana, la registraba y también anotaba la alineación de la cabeza de la lechuza.

Las pruebas de orientación de la cabeza han aportado información abundante sobre cómo determina el origen de un sonido. Un dato muy importante del proceso es que puede conseguir una precisión máxima incluso con sonidos que terminan antes de que se inicie el movimiento de cabeza. Ello indica que su sistema auditivo fija el acimut y la elevación de un sonido sin mover la testa y utiliza después la información para dirigir la respuesta de orientación de la cabeza. Las pruebas de movimiento de la cabeza muestran asimismo que la precisión se reduce a medida que aumenta el ángulo entre el origen del sonido y la orientación de la cabeza.

Nuestros experimentos y los de otros han demostrado que la capacidad de la lechuza para identificar el origen de un sonido depende de la presencia en éste de frecuencias altas. Aunque posee un oído sensible a un amplio espectro de frecuencias, desde 100 hertz (ciclos por segundo) hasta 12.000 hertz, puede localizar con precisión únicamente aquellos sonidos cuyas frecuencias se hallan entre 3000 y 9000 hertz. Además, los ensayos donde se le tapaba un oído ponen de manifiesto que necesita los dos para la localización precisa de los blancos. Si se le bloquea uno, comete grandes errores en la dirección de éste.

Con estos datos por delante, nos aprestamos a investigar cuál era la información exacta que la lechuza extraía de los sonidos naturales. Para localizar la fuente emisora, debe determinar la dirección de propagación de las ondas sonoras según la información que recogen los detectores situados en dos puntos, a saber, sus oídos. La información espacial más útil se obtiene al comparar información procedente de estos dos puntos, puesto que las diferencias entre ellos no dependen del nivel sonoro absoluto, sino de la orientación de los oídos en el campo sonoro.

Un dato valioso es el desfase temporal en la llegada del sonido a los órganos de la audición. Cuando el sonido procede directamente de un lado, la diferencia es máxima; cuando el sonido se halla directamente frente al ave, no hay diferencia en el tiempo de llegada a los dos oídos. Entre estos límites, el intervalo de desfase varía con el ángulo del sonido en el plano horizontal. De ahí que la demora de tiempo pueda

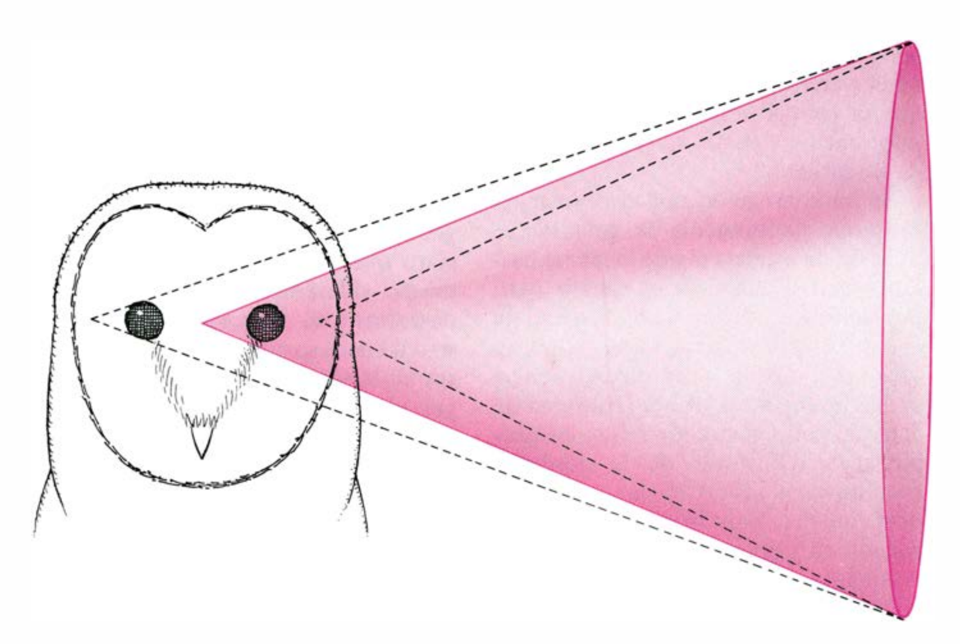


SISTEMA BIPOLAR DE COORDENADAS que sirve para cartografiar la posición de la cabeza en pruebas de orientación. Cada dirección viene especificada por dos medidas: el ángulo con respecto a la horizontal y el ángulo con respecto a la vertical, en relación al cual la lechuza debe girar la testa para mirar en aquella dirección. Mirando hacia adelante, que es como normalmente se halla posada el ave, la orientación es de 0 grados (horizontal) y 0 grados (vertical). El mapa abarca los 180 grados del espacio situado frente al ave. Los mapas que vienen en este artículo utilizan el mismo sistema de coordenadas.

proporcionar información acerca del acimut del sonido. Lo que no le basta para situar con exactitud el sonido, pues hay ruidos que proceden de direcciones distintas y sin embargo originan la misma diferencia de tiempo. En el espacio tridimensional, estas direcciones forman un cono alrededor del eje que une los oídos de la lechuza.

Valga el caso como ejemplo de lo inadecuado de una única señal, la que sea, para proporcionar información acerca de los ángulos horizontal y vertical de la fuente de un sonido. Para especificar una localización en ambas dimensiones se precisan dos señales independientes. En el caso de la lechuza, la información adicional la proporcionan las diferencias en la sensibilidad direccional de los órganos de la audición.

La sensibilidad direccional la aporta la gola facial. Como una mano que se ahueca detrás de la oreja, los conductos de la gorguera amplifican el sonido y refuerzan la sensibilidad del oído a los sonidos que proceden de determinadas direcciones. La intensidad de amplificación y de sensibilidad direccional que proporcionan las plumas de la gorguera facial varía, de manera espectacular, con la frecuencia del sonido. Ello se debe a una de las propiedades de las mismas ondas sonoras; cuando éstas, en su avance, encuentran un objeto, lo rodean o se reflejan en él. Que ocurra lo primero o lo segundo dependerá de la longitud de onda del sonido y del tamaño del objeto. Si la longitud de onda es larga comparada con el objeto, las ondas tienden a propagarse bordeándolo; pero si es corta, las ondas rebotan, por reflexión, en la misma dirección.



CONO DE CONFUSION (*color*) formado por aquellas direcciones que la lechuza no puede distinguir entre sí sirviéndose sólo de la demora de tiempo. El retardo temporal proporciona información acerca del ángulo horizontal de un manantial sonoro. Puesto que los oídos izquierdo y derecho suelen hallarse a distancias ligeramente distintas del origen de un sonido, las ondas llegan a ellos con un ligero intervalo de desfase. Cuanto mayor es el ángulo de una fuente sonora con respecto al plano frontal del ave, mayor es la demora de tiempo. Sin embargo, hay muchas direcciones que originan la misma longitud de trayectoria (*líneas discontinuas*) y, por tanto, la misma demora de tiempo; el ave habrá de servirse de datos adicionales para distinguir una dirección de otra. En el espacio tridimensional, estas direcciones forman un cono cuyo vértice se halla entre los dos órganos de la audición. Los dibujos de este artículo son de Tom Prentiss.

En virtud de ese fenómeno, las frecuencias inferiores a 3000 hertz no las refleja de un modo satisfactorio la gorguera. Puesto que la acción de encauzamiento del oído depende de su capacidad para reflejar sonidos, su sensibilidad direccional a frecuencias bajas es relativamente pobre. Tenemos así que, a 3000 hertz, el oído izquierdo es sólo ligeramente más sensible a los sonidos procedentes de una zona situada entre los 20 y los 40 grados a la izquierda que a los sonidos que llegan de otras direcciones. El oído derecho goza de un grado de sensibilidad semejante a la derecha. Habida cuenta de que la sensibilidad de cada oído a las frecuencias bajas sólo cambia de forma gradual con la dirección, la comparación de intensidades sonoras a frecuencias bajas únicamente puede proporcionar una señal espacial grosera. Por si fuera poco, esa diferencia no revela nada sobre la elevación del sonido.

Otra cosa ocurre con ondas sonoras de frecuencia superior. Los órganos de la audición muestran una sensibilidad mucho más fina a la dirección del sonido; una ligera modificación en la dirección del sonido produce un gran cambio en la intensidad percibida. Además, en lugar de ser más sensibles a la derecha o a la izquierda, el oído derecho es más sensible por encima del plano horizontal y el izquierdo lo es por debajo de dicho plano. Esta sensibilidad responde a una simetría anatómica insólita. La golilla de la izquierda se dirige ligeramente hacia abajo, y la abertura del oído y el halda preaural se hallan situadas en una posición más alta en la golilla del lado izquierdo. En el lado derecho ocurre lo contrario. En consecuencia, conforme la fuente sonora se desplaza hacia arriba, las componentes de alta frecuencia de un sonido natural se hacen más intensas en el oído derecho y más moderadas en el izquierdo. Al descender el manantial sonoro, los sonidos se hacen más intensos en el oído izquierdo. Puesto que a altas frecuencias la intensidad percibida cambia rápidamente con la elevación, esta señ?l ofrece información muy precisa.

Se trata de una información valiosa y compleja a un tiempo. La magnitud de la diferencia en intensidad varía según la frecuencia del sonido, debido a la mayor capacidad de la gorguera para reflejar sonidos de frecuencias más elevadas. La dirección indicada varía asimismo con la frecuencia, puesto que la localización horizontal viene suministrada por las frecuencias bajas y, por las altas, la vertical. De ahí que el sistema auditivo de la lechuza tenga que comparar las intensidades detectadas en cada órgano para cada frecuencia. Se denomina espectro interaural a la comparación de intensidades hecha frecuencia a frecuencia.

Ya que los sonidos de baja frecuencia proporcionan señales para el acimut y los de alta frecuencia señales para la elevación, el espectro interaural podría aportar, por sí solo, información suficiente para que la lechuza descubriese la presa. Hay muchas pruebas que apoyan la hipótesis de que se sirven de dicho espectro. La precisión de la arremetida del ave aumenta de manera clara cuando crece la amplitud de banda (el número de frecuencias que contiene un sonido). A partir de diferencias en la intensidad de un único tono (un sonido de frecuencia única), la lechuza puede determinar la dirección en una sola dimensión; a medida que el espectro se amplía se dispone de más diferencias de intensidad, y sus valores indican el ángulo de la fuente en más de un plano. Los experimentos en que se tapa uno de los oídos demuestran, sin tanto rodeo, que compara intensidades. Una lechuza entrenada a la que se obstruye su órgano derecho arremete a la izquierda sin llegar al blanco; con el oído izquierdo tapado arremete a la derecha más allá del blanco. Puesto que

los errores tienen a la vez componentes de elevación y de acimut, cabe suponer que el animal obtiene ambos tipos de información a partir de comparaciones de intensidad.

Se ha conseguido una confirmación posterior extirpándole la gola facial. Sin ésta, la lechuza puede localizar bastante bien el acimut de un sonido, pero no identifica su ubicación vertical: el ave se orienta siempre hacia un punto sobre la horizontal, con independencia de la elevación del blanco. Ello está de acuerdo con nuestra hipótesis: las propiedades reflectoras de la gorguera constituyen el soporte de la sensibilidad direccional a altas frecuencias, que le permite identificar el ángulo vertical de la fuente. La remoción de la gorguera elimina la capacidad de discriminar entre elevaciones. Se conserva una cierta disparidad en las intensidades, porque las aberturas auditivas y las haldas preaurales se hallan situadas asimétricamente en el pliegue de piel que sostiene las plumas de la gorguera, pero ello no basta para identificar la elevación del sonido.

De este modo, a partir de los resultados obtenidos en las pruebas de orientación de la cabeza, se ha confirmado ampliamente que las lechuzas se basan en el espectro interaural para ubicar sonidos en acimut y elevación. Sin embargo, a partir de otros hallazgos resulta claro que, en parte, el ave utiliza asimismo diferencias de sincronía para descubrir a su presa invisible. El retraso o demora se manifiesta en dos aspectos de la señal binaural. En primer lugar, el sonido empieza y acaba antes en el oído más cercano al origen; la localización temporal de las principales discontinuidades de intensidad del sonido es asimismo ligeramente distinta en cada órgano de la audición. Estas diferencias reciben la denominación global de disparidad transitoria. En segundo lugar, a lo largo de toda la duración del sonido, las ondas que llegan al oído más alejado se hallarán ligeramente retrasadas. Con una única frecuencia, esta diferencia en la sincronía de las formas de onda se conoce por demora de fase; en sonidos naturales más complejos, constituidos por muchas frecuencias, se denomina disparidad temporal progresiva.

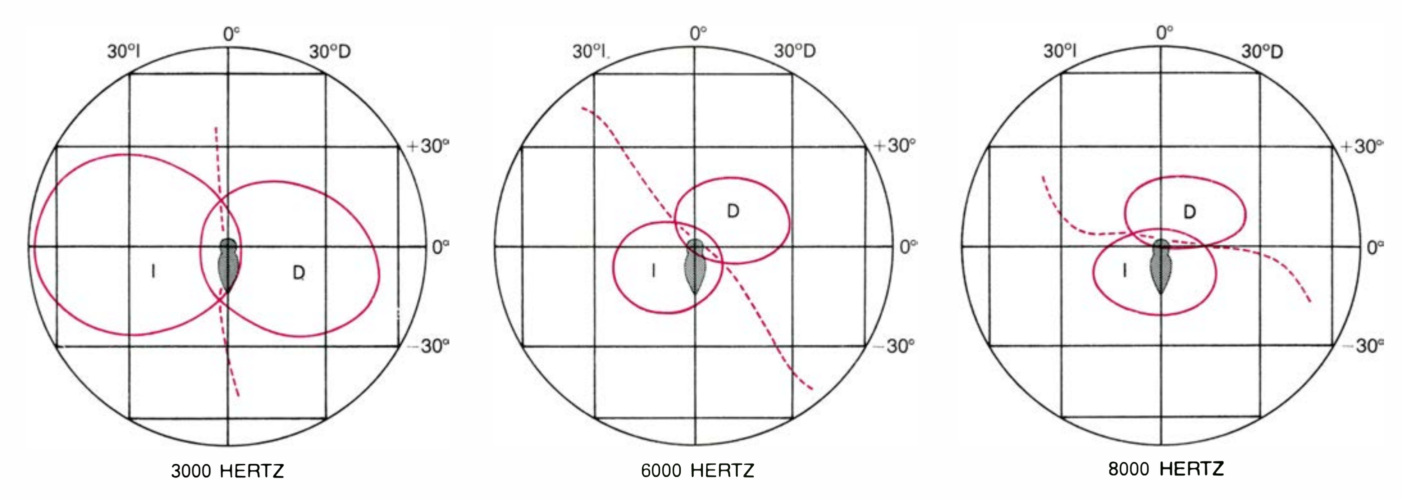
En la naturaleza, las disparidades progresiva y transitoria vienen a coincidir en magnitud; varían con los cambios en el acimut de la fuente sonora. Sin embargo, presentan distintas ventajas a la hora de localizar los sonidos. La disparidad temporal progresiva puede medirse repetidamente mientras perdura el sonido. En cambio, la disparidad transitoria sólo puede verificarse de forma intermitente, pero es más difícil que se confunda con ecos.

Los humanos nos fundamos en la disparidad transitoria y en la progresiva para ubicar la fuente sonora. Las lechuzas parecen emplear sólo la diferencia progresiva. Al igual que otros tipos de información espacial, la disparidad progresiva presenta una importante ambigüedad. Se comprenderá mejor para el caso de un tono que proceda directamente de un lado. Las señales detectadas por el ave son ondas sinusoidales (regulares); debido a la diferencia de distancias entre el manantial sonoro y uno y otro órgano de la audición, las ondas se hallarán ligeramente desfasadas entre sí. La magnitud del retraso de fase así creado dependerá, a la vez, de la frecuencia del tono y de la diferencia entre los oídos. Conforme aumenta la frecuencia de la onda o la distancia entre los oídos, la onda pasa a través de una parte mayor de su ciclo mientras rodea la cabeza hasta alcanzar el otro oído; de ahí que la demora de fase sea mayor.

Cuando la frecuencia del tono es tan alta que la onda pasa a través de exactamente la mitad de su ciclo antes de alcanzar el oído más alejado, el desfase corresponde a la mitad de la longitud de onda. Un retraso de este tipo podría estar causado por un sonido que procediera directamente de la izquierda o directamente de la derecha de la lechuza, puesto que la diferencia en longitud de recorrido es la misma para ambas direcciones. Por ello, es imposible que el sistema auditivo determine la dirección del sonido únicamente sobre la base de la disparidad progresiva. A frecuencias superiores, la situación empeora. Cuando la longitud de onda es igual a la distancia entre los oídos, por ejemplo, no hay retraso de fase, puesto que la onda viaja a lo largo de todo su ciclo mientras rodea la cabeza. Esta relación podría corresponder a un sonido que viniera de la derecha, de la izquierda o de enfrente; si se sirviera sólo del retraso de fase, el ave no tendría forma de determinar cuál es el caso.

Las longitudes de onda a las que apa-

REGISTRO SONORO, que muestra dos tipos de diferencias en la sincronía del sonido entre los oídos izquierdo y derecho. En este esquema, la altura del registro indica la presión causada por las ondas sonoras en cada oído. Cuando un sonido procede de la derecha, la onda empieza y termina antes en el órgano derecho; también allí tienen lugar, ligeramente antes, cambios importantes de intensidad. Esas diferencias reciben la denominación global de disparidad transitoria. Además, a lo largo de la duración del sonido, las formas de la onda se hallan ligeramente más adelantadas en el oído derecho: se trata del fenómeno de disparidad progresiva. El hombre se basa en ambas diferencias para situar el origen de un sonido. La lechuza utiliza la disparidad progresiva, pero no parece que aproveche la transitoria.



SENSIBILIDAD DIRECCIONAL de los oídos de la lechuza, que cambia con la frecuencia de las ondas sonoras. La acción concentradora de los conductos de la gorguera facial resulta en un aumento de la sensibilidad de los oídos a los sonidos procedentes de determinadas direcciones frente a los procedentes de otras. A frecuencias bajas, la zona de sensibilidad máxima del oído izquierdo se halla a la izquierda, y la del oído derecho, a la derecha. La gorguera refleja mejor las ondas de altas frecuencias. Puesto que el conducto izquierdo está desplazado hacia abajo y el derecho hacia arriba, al ir aumentando la frecuencia, las zonas de mayor sensibilidad adoptan una disposición vertical: el oído derecho es más sensible hacia arriba y, hacia abajo, el izquierdo. A 8000 hertz (ciclos por segundo) las regiones de audición más sensible se sitúan casi directamente por encima y por debajo del ave. El sonido natural suele tener varias frecuencias; la lechuza aprovecha así las diferencias de intensidad para identificar tanto la dirección horizontal como la vertical de una fuente sonora.

recen estas ambigüedades dependen de la distancia entre los oídos. Los de la lechuza están separados unos cinco centímetros; por ello, la ambigüedad de fase surgirá a una longitud de onda de 10 centímetros o menos, lo que corresponde a frecuencias de 3000 hertz o superiores. Puesto que no tiene dificultad en determinar el acimut, incluso de tonos de alta frecuencia, Konishi y yo supusimos que a frecuencias altas el ave debía fiar en alguna fuente de información distinta de la disparidad temporal progresiva. Un candidato plausible para aportar los datos adicionales era la disparidad transitoria, que no resulta afectada por los cambios de frecuencia; se sabe que otras especies, el hombre incluido, dependen de ella a altas frecuencias.

Andrew Moiseff y Konishi, en un experimento ulterior de orientación de cabeza, han demostrado de manera concluyente lo erróneo de esta deducción. Presentaron sonidos directa e independientemente a ambos oídos, mediante pequeños altavoces implantados en los canales auditivos. Esta técnica les permitió eliminar disparidades transitorias y diferencias de intensidad entre ambos órganos, al tiempo que variaban la disparidad temporal progresiva. La demora entre las formas ondulatorias del sonido en los dos oídos podía ajustarse en fracciones muy pequeñas, de hasta un microsegundo. En respuesta a disparidades progresivas de sólo 10 microsegundos o de hasta 80 microsegundos, la lechuza daba rápidos giros horizontales de cabeza, que correspondían aproximadamente al ángulo implicado en la diferencia en curso. Esta respuesta sugiere, a falta de otras señales interaurales, que la lechuza continúa utilizando el retraso de fase, o disparidad progresiva, a frecuencias altas.

El descubrimiento era sorprendente, pues implicaba que la lechuza disponía de algún sistema para superar la ambigüedad de fase. Además, para que el sistema auditivo perciba esas pequeñas diferencias en la sincronización de las formas ondulatorias en los dos oídos, debe recibir información específica acerca de señales acústicas que se presentan a 7000 hertz, o sea, una vez cada 143 microsegundos. Lo que merece destacarse, habida cuenta de que los impulsos nerviosos que conducen esa información de la cóclea al cerebro duran más de 1000 microsegundos.

La disparidad temporal progresiva es de gran utilidad para la lechuza. Mas, por sí sola, la señal no tiene la precisión que ella necesita para cazar. Con altavoces implantados en los canales auditivos, respondía a las distintas disparidades temporales progresivas con giros que variaban en magnitud hasta 15 grados en cada dirección. En cambio, la mayor desviación típica del error para lechuzas que respondían a blancos externos era de sólo 2,5 grados en cada dirección; el error no solía alcanzar los 1,5 grados. Por tanto, debían combinarse otras señales con la disparidad progresiva. Se sabía que las diferencias de intensidad ayudan a especificar la localización, pero intentamos asimismo descubrir si el ave utilizaba disparidades transitorias.

Nuestra primera hipótesis, según la cual la lechuza se basaba en la disparidad transitoria para compensar la ambigüedad de fase a frecuencias elevadas, se ha visto de plano invalidada. Para fijar el acimut de un tono, la lechuza no se apoya en absoluto en las disparidades transitorias. En pruebas de orientación de la cabeza se le presentaban tonos de 7000 y 8000 hertz. El altavoz se hallaba a veces a 30 grados a la derecha y otras a 30 grados a la izquierda. En estas situaciones se observó un curioso comportamiento: con el blanco a 30 grados a la derecha, la lechuza giraba a veces 30 grados a la izquierda, y viceversa.

Esta confusión constituye una clara manifestación de ambigüedad de fase. El período cíclico de un tono de 7000 hertz es de 143 microsegundos; el de un tono de 8000 hertz, 125 microsegundos. Con una fuente sonora a 30 grados a la derecha o a la izquierda, estos tonos recorren aproximadamente medio ciclo al superar la distancia que separa los oídos. De ahí que un tono que proceda de 30 grados a la derecha proporcione el mismo retraso de fase que otro que llega desde 30 grados a la izquierda. La ambigüedad se presenta sólo en la disparidad progresiva; si la lechuza se hubiera basado en la disparidad transitoria habría adivinado inmediatamente la localización correcta.

Este resultado confirma que el ave continúa apoyándose en información de demora de fase, incluso a frecuencias elevadas, a pesar de su ambigüedad, pero no parece que se funde en la disparidad transitoria. Puesto que la longitud de onda de los sonidos de baja frecuencia es considerablemente mayor que la distancia entre uno y otro órgano de la audición, a frecuencias bajas no existe ambigüedad de fase. Además, la presencia de varias frecuencias en los

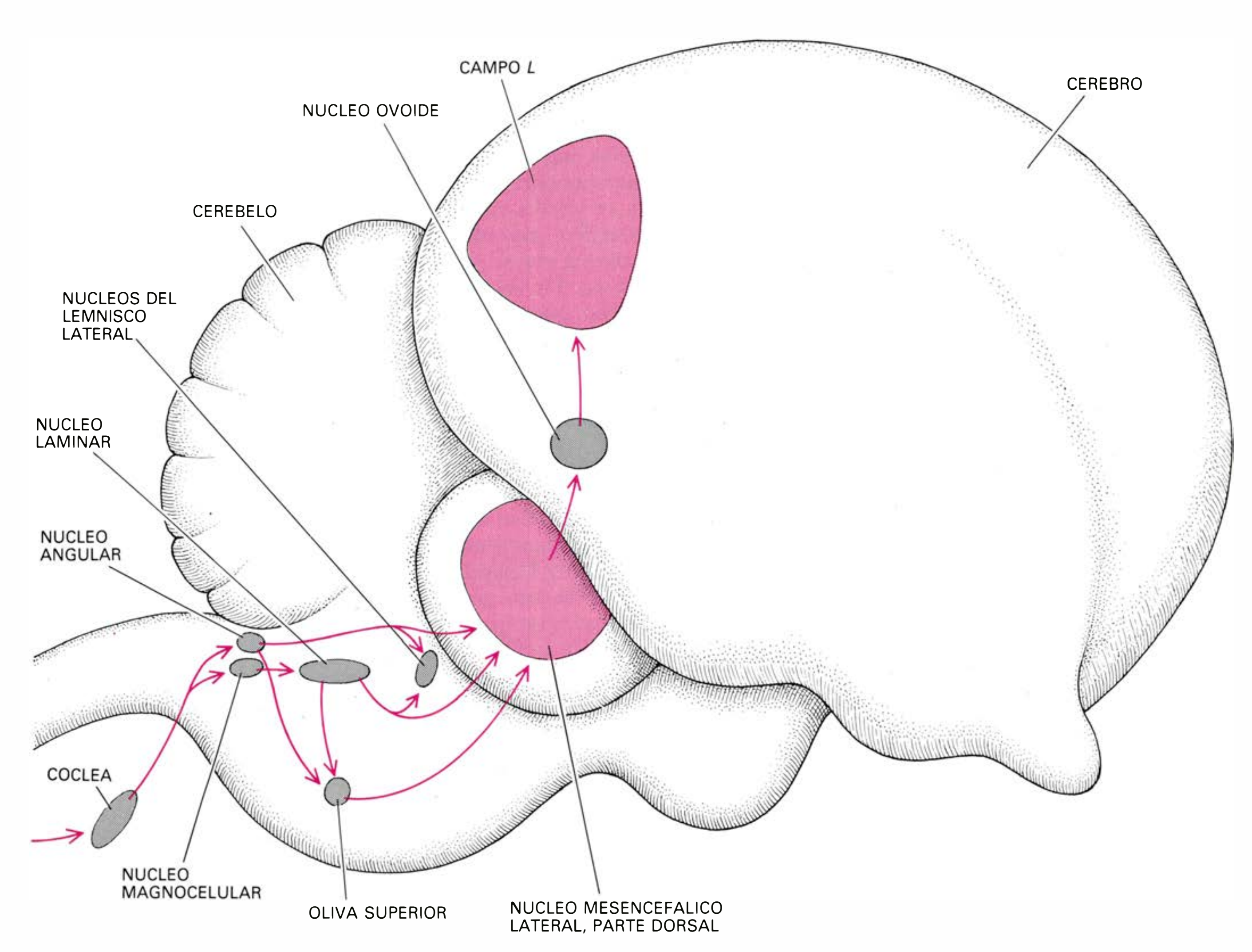
sonidos naturales coadyuva a la resolución de las ambigüedades de fase. Cuando surgen muchos pares de direcciones distintos, a partir de frecuencias asimismo distintas, selecciona la dirección en el espacio que encaja con un miembro de cada par.

Las pruebas de orientación de la cabeza muestran claramente que la lechuza se basa en dos tipos de información para determinar el origen de un sonido: el espectro interaural y la diferencia temporal progresiva. Esta última informa acerca del acimut de la fuente sonora. Las diferencias de intensidad entre los oídos proporcionan, a la vez, señales para el acimut y la elevación. Averiguar de qué manera el sistema auditivo transforma estas señales en una imagen neuronal de localización de sonido requiere un enfoque experimental distinto.

La técnica que utilicé, en primer lugar con Konishi y posteriormente en experimentos independientes, consistía en insertar un microelectrodo en el cerebro de una lechuza anestesiada y buscar así puntos de actividad neuronal mientras se la estimulaba con ruidos. Konishi y yo abordamos los experimentos con el mismo aparato empleado en las pruebas de orientación de la cabeza. Una vez anestesiada la lechuza, manteníamos rígida su cabeza en un armazón estereotáxico especial. Se introducía entonces un microelectrodo en el cerebro hasta que se registraban impulsos nerviosos procedentes de una sola neurona. Moviendo el altavoz diana alrededor de la lechuza se cartografiaban las regiones del espacio a las que respondía la neurona.

Asignar el origen de un sonido en el espacio requiere un considerable procesamiento neuronal. Sabido esto, comenzamos nuestro estudio explorando estructuras bastante alejadas del trayecto auditivo: en el cerebro medio y anterior. El principal centro auditivo del mesencéfalo de las aves es la llamada parte dorsal del núcleo mesencefálico lateral (*nucleus mesencephalicus lateralis pars dorsalis*, MLD), que corresponde a la estructura que en el cerebro de los mamíferos se denomina colículo inferior. Los impulsos nerviosos que llegan a ese centro ya han sido procesados en uno o más núcleos situados en un segmento anterior del trayecto auditivo. Un segmento posterior de ese trayecto, una zona denominada campo *L*, es el centro receptor primario del prosencéfalo para los impulsos auditivos. (Esta estructura corresponde al córtex auditivo de los mamíferos; las aves carecen de un análogo exacto del córtex auditivo.)

Tanto en el MLD como en el campo *L*, la gran mayoría de las neuronas no responden de manera precisa a señales espaciales. Algunas neuronas presentan su nivel de actividad más elevado en respuesta a sonidos procedentes de una determinada región del espacio,



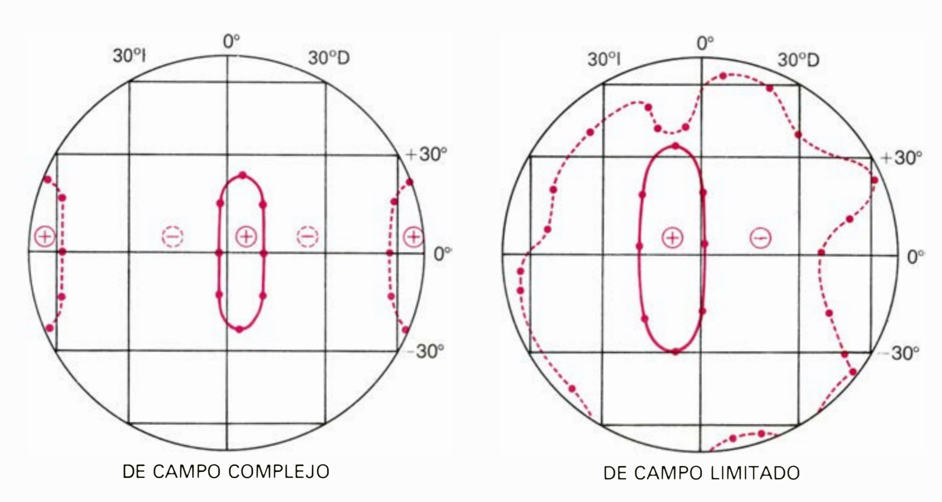
TRAYECTO AUDITIVO del sistema nervioso central de la lechuza; conduce de la cóclea a una zona del cerebro anterior denominada campo *L*. A lo largo de este trayecto, los impulsos nerviosos sufren un procesamiento bastante complejo: los datos sobre la sincronía y la frecuencia del sonido se convierten en información acerca de la localización de la fuente sonora. Buena parte de este procesamiento acontece en varios centros del cerebro inferior. Cuando los impulsos alcanzan el núcleo del cerebro medio denominado núcleo mesencefálico lateral parte dorsal (MLD), se dirigen a una red de neuronas que responden a los sonidos procedentes de zonas específicas; la distribución de estas zonas dibuja una representación bidimensional del espacio situado frente al ave. La información sobre la localización exacta del sonido pasa seguidamente al campo *L*, que corresponde al córtex auditivo de los mamíferos.

pero las fronteras de la región no son precisas, y varían mucho con la intensidad del ruido. Otras neuronas se muestran todavía menos específicas en su respuesta: se excitan en presencia de sonidos procedentes de prácticamente todas las direcciones. Es probable que la contribución de estas neuronas no sea establecer la ubicación sino detectar o identificar los sonidos.

Sin embargo, hay dos tipos de neuronas, en el MLD y en el campo *L*, altamente sensibles a la localización de sonidos. La primera de ellas, la denominada neurona de campo complejo, se encuentra en gran número en el MLD, por lo general en grupos dispersos entre otros tipos de neuronas. La actividad de las neuronas de campo complejo viene estimulada por sonidos provocados en varias regiones distintas del espacio, los llamados campos excitadores. Las neuronas se excitan mucho menos, si es que no llegan a inhibirse, al responder a sonidos que surgen de las regiones del espacio situadas entre los campos excitadores.

Cuando se calcula para una única neurona de campo complejo la localización de los centros de los campos excitadores en el espacio, se observa una interesante correspondencia. Los campos excitadores son los mismos que las regiones del espacio que la lechuza confunde en condiciones de ambigüedad de fase. Los múltiples campos excitadores representan las regiones desde donde los sonidos que alcanzan los oídos generarán retrasos de fase de magnitud equivalente. Por tanto, cada neurona de campo complejo parece ser sensible a un determinado retraso de fase, a una determinada frecuencia. La presencia de varios campos excitadores para cada neurona constituiría el recíproco físico de la confusión espacial debida a la ambigüedad de fase.

El segundo tipo de neurona, la denominada neurona de campo limitado, que se encuentra tanto en el MLD como en el campo *L*, responde de una manera todavía más específica. Las células de campo limitado sólo resultan excitadas por sonidos que proceden de una región del espacio. Las regiones a las que responden las neuronas de campo limitado son típicamente elípticas. Su tamaño varía en acimut de siete a 42 grados, y en elevación desde los 23 grados hasta una banda completa que se extiende frente al ave. A diferencia de otras neuronas del trayecto auditivo, las de campo limitado son extremadamente selectivas, respondiendo sólo a cambios en la localización; las variaciones importantes de la intensidad sonora provocan pocas alteraciones, o ninguna, en los límites precisos de la región receptiva.



DOS TIPOS DE NEURONAS del cerebro de la lechuza, sensibles a sonidos procedentes de direcciones específicas. Se han representado aquí los campos receptivos de los dos tipos celulares. Los campos receptivos son las regiones del espacio en cuyo interior los sonidos producen una respuesta excitadora de la neurona. Las neuronas de campo complejo, que se encuentran en el MLD, tienen varias de estas zonas. A la izquierda se muestra una pauta receptora típica de campo complejo. Las zonas receptivas corresponden a las direcciones que originan disparidades progresivas idénticas en la sincronía de las ondas sonoras. Entre estas regiones, los sonidos desencadenan una respuesta inhibidora. Las neuronas de campo limitado se encuentran tanto en el MLD como en el campo *L*. A la derecha se muestra una pauta receptora típica de campo limitado. Las células de campo limitado tienen una única zona receptora. Los sonidos procedentes del exterior de esta zona producen una fuerte respuesta inhibidora. Cada célula de campo limitado responde a una determinada diferencia en sincronía sonora e intensidad entre los órganos auditivos.

Parte de la nitidez de los bordes de las regiones receptivas de las neuronas de campo limitado se debe a que los sonidos que provienen del exterior de la región excitadora inhiben la respuesta de la célula. El efecto inhibidor se hace más fuerte a medida que la localización del sonido se acerca a la frontera de la región excitadora, en cuyo punto el efecto inhibidor cambia para tornarse excitador. Aunque las regiones espaciales que originan una respuesta excitadora son bastante grandes, en su interior se encuentra una región más pequeña que corresponde a la mejor respuesta. Estas "regiones óptimas" varían en tamaño de 2,5 a 15 grados, en acimut, y de cinco a 45 en elevación.

Tanto en el MLD como en el campo *L* se encuentran unidades de campo limitado. Pero su distribución en los dos centros difiere sustancialmente. En *L*, las neuronas de campo limitado constituyen sólo alrededor del 15 por ciento de la población total de células nerviosas, y con frecuencia se hallan dispersas entre otros tipos celulares. En cambio, en el MLD estas células se concentran en los bordes laterales y frontal del núcleo, entremezcladas únicamente con unas cuantas neuronas del tipo de campo complejo. Tras explorar y cartografiar de forma intensiva el MLD, se hizo evidente que la disposición de las neuronas de campo limitado en este núcleo constituía un mapa del espacio bidimensional, con la distribución de las zonas receptivas de las neuronas resiguiendo los contornos espaciales.

Sin embargo, el mapa está distorsionado: la zona que se halla frente al ave, la región de máxima agudeza auditiva, está representada desproporcionadamente. Los acimuts sonoros se disponen en el plano horizontal. En el lado derecho de la estructura que representa el mapa se encuentran neuronas excitadas por sonidos que se originan entre los 15 grados a la derecha y los 60 grados a la izquierda. En el lado opuesto hay neuronas estimuladas por sonidos procedentes desde los 15 grados a la izquierda hasta los 60 grados a la derecha. Esta disposición significa que los 30 grados de espacio situados frente al ave se hallan representados en ambos lados del mapa y, por tanto, por dos conjuntos de neuronas. Además, el mapa se dispone de manera que en ambos lados la población de neuronas que representa los 30 grados situados al frente es desproporcionadamente grande. La representación doble y el gran número de neuronas a cada lado suponen que los 30 grados frontales se analizan con gran precisión, hecho que puede explicar la particular agudeza de la lechuza a la hora de localizar sonidos en esta zona.

Las elevaciones sonoras se disponen transversalmente sobre el mapa. Las "regiones óptimas" de las neuronas de campo limitado van de los 40 grados hacia arriba a los 80 grados hacia abajo. Los campos superiores se hallan en la cúspide de la superficie curvada del mapa, y los inferiores en la parte baja.

¿Cómo construye tan notable mapa el sistema nervioso? Los experimentos de Moiseff y Konishi han contribuido de manera sustancial a la comprensión del proceso. En estas pruebas se colocaban altavoces en los canales auditivos de lechuzas anestesiadas. Cuando se presentaba un sonido por separado a los dos oídos, se observaban las diferencias de sincronía y de intensidad necesarias para disparar una respuesta en cada neurona. Después, estos valores se correlacionaban con las zonas espaciales que se sabía que excitaban cada conjunto de neuronas. Los resultados de las investigaciones muestran que las unidades de campo limitado en el mapa son bastante sensibles a diferencias progresivas en la sincronía. Estas observaciones se corresponden perfectamente con la dependencia del ave respecto de las disparidades progresivas, demostrada en los experimentos de orientación de la cabeza.

MAPA BIDIMENSIONAL del espacio frontal, tal como se encuentra en el MLD de la lechuza. El mapa está constituido por neuronas de campo limitado. Cada célula responde a sonidos de una región específica del espacio. La disposición de las zonas receptoras de las neuronas sigue contornos espaciales. Arriba se muestra la región del lado izquierdo del cerebro que responde a sonidos procedentes de la derecha. Las neuronas de esta región son sensibles a sonidos de una zona que va desde los 15 grados a la izquierda hasta los 60 grados a la derecha, y desde los 40 grados arriba hasta los 80 grados abajo. De este modo, los registros de electrodos a través de la estructura corresponderán a una serie de regiones espaciales dispuestas verticalmente; se indican tres de esos registros y las regiones espaciales a las que corresponden. En una estructura análoga, situada en el lado derecho del cerebro, se hallan las neuronas que responden a sonidos que proceden de la izquierda. La región que va de los 15 grados a la izquierda a los 15 grados a la derecha se halla así representada en ambos lados del cerebro y aparece en el mapa con un número de neuronas desproporcionadamente grande. Aquí reside la zona de mayor agudeza para localizar una fuente sonora.

Las neuronas de campo limitado responden sólo a una banda extremadamente estrecha de demoras progresivas de tiempo: la magnitud de la banda va de los 40 a los 100 microsegundos. Incluso dentro de este minúsculo espectro, una demora determinada disparaba siempre la mayor respuesta. Cambiando la diferencia progresiva en sólo 10 microsegundos, la intensidad de la respuesta de la neurona podía variar hasta en un 75 por ciento. Este grado de sensibilidad complementa (y ayuda a explicar) la precisión de la lechuza en la caza aérea.

La disparidad progresiva que desencadena la mayor respuesta pertenece asimismo a la región espacial a la que la célula es sensible. Las neuronas que responden a sonidos que proceden del frente resultan excitadas de modo máximo por disparidades progresivas pequeñas. Las neuronas que responden a sonidos de ángulos mayores requieren disparidades mayores. Estos resultados confirman, a nivel celular, la conclusión obtenida a partir de las pruebas de comportamiento, a saber: la lechuza se basa, sobre todo, en la diferencia progresiva de la sincronía entre los oídos para determinar la localización en el plano horizontal.

Tales experimentos han contribuido a confirmar y explicar por qué la lechuza fía en diferencias de intensidad sonora. Variando la intensidad relativa de los sonidos ofrecidos, de manera independiente, a los oídos de una lechuza anestesiada se encontró que, por cada neurona de campo limitado, había una diferencia de intensidad que evocaba la respuesta máxima. Al alejarse de este valor, la respuesta disminuía y, finalmente, cesaba. La pauta no resultaba afectada por los cambios en el nivel sonoro medio; sólo dependía de la diferencia de intensidad entre los oídos.

Los resultados obtenidos al sondar el cerebro muestran que el mapa espacial en el MLD responde a las mismas señales que el ave utilizaba para situar el acimut y la elevación de un sonido: las disparidades temporales progresivas y el espectro interaural. Mediante una transformación exquisitamente precisa, el sistema auditivo convierte estas señales en información espacial. La disposición de las células en el mapa implica que las neuronas vecinas responden a señales que sólo son ligeramente distintas. Además, el orden de los campos excitadores en el mapa sigue la continuidad espacial. ¿Cómo establece el cerebro de la lechuza estas conexiones precisas?

Fibrinógeno y fibrina

La proteína plasmática fibrinógeno se convierte en fibrina formando un coágulo sanguíneo. Con el tiempo, el coágulo se desintegra. Ambos procesos se han desentrañado ya merced al descriframiento total de su estructura molecular

Russell F. Doolittle

Nos pinchan y sangramos, pero la hemorragia se detiene; la sangre coagula. Los fragmentos celulares denominados plaquetas tienen la propiedad de adherirse al punto donde existe la solución de continuidad, cerrando parcialmente la abertura. Las plaquetas poseen otra importante función: en su superficie se concita todo un ejército enzimático cuyo producto final será una proteasa: la trombina. El blanco de tal proteasa es el fibrinógeno, molécula de gran tamaño que circula disuelta en el plasma. Bajo la acción de la trombina, el fibrinógeno experimenta una notable transformación: sus moléculas alteradas se encadenan espontáneamente formando un polímero fibrilar, denominado fibrina, principal constituyente del coágulo sanguíneo. Los haces de fibrina se organizan en una red; este trenzado, que surge sólo en las proximidades de las plaquetas adheridas, donde se genera la trombina, transforma localmente el plasma sanguíneo en un gel.

¿Qué rara especie de molécula es el fibrinógeno, que después de sufrir el ataque de una proteasa (la trombina) polimeriza espontáneamente? Aplacemos la respuesta para después de haber considerado otro fenómeno de la coagulación. El coágulo de fibrina no está diseñado para persistir indefinidamente. Se trata de un remedio transitorio que se establece rápidamente y que muy pronto se sustituirá por otro más definitivo. Hay otras razones, quizá más apremiantes, que obligan a la disolución del coágulo. Un pequeño fragmento de un coágulo (un émbolo) desprendido en el sistema circulatorio puede ocluir un vaso de pequeño calibre; un embolismo en el pulmón puede ser fatal. Un trombo, es decir, un coágulo formado inoportunamente como consecuencia de un proceso patológico, como la aterosclerosis, y no en respuesta a una lesión del vaso, puede desencadenar un infarto o una apoplejía.

Para minimizar los peligros del embolismo y reducir la posibilidad de una trombosis, así como para destruir un coágulo normal, en el proceso de curación de las heridas, se requiere un sistema proteolítico complementario del que genera trombina. Dicho sistema da lugar al enzima plasmina, que destruye la red de fibrina y restaura la fluidez normal del plasma. Habrá que completar la formulación de nuestra pregunta anterior: ¿qué rara especie de molécula es el fibrinógeno, que puede polimerizarse por el ataque de una proteasa (la trombina) y luego despolimerizarse al ser atacada por otra proteasa (la plasmina)? Esto es, ¿cómo se explica la función del fibrinógeno y la fibrina a partir de su estructura? La respuesta constituye la razón de este trabajo.

La estructura del fibrinógeno ha sido objeto de acalorado debate a lo largo de los últimos 10 o 15 años. Si bien aún se asiste a polémicas residuales entre algunos investigadores de la química de las proteínas, disponemos ya de múltiples datos que apoyan el tipo de estructura que expondremos a continuación.

Una de las fuentes de controversia la constituyen los datos que aporta la microscopía electrónica. En 1959, Cecil E. Hall y Henry S. Slayter, del Instituto de Tecnología de Massachusetts, publicaron unas excelentes micrografías que sugerían claramente una estructura trinodular del fibrinógeno. En aquel tiempo, los químicos estructurales iban bastante rezagados de los microscopistas y, si bien la forma trinodular se aceptó mayoritariamente durante algún tiempo, no recibió el soporte de los químicos. La metodología del microscopio electrónico cambió rápidamente a comienzos de la década de los sesenta. Hall y Slayter habían obtenido sus micrografías con la técnica del sombreado, en tanto que los microscopistas preferirían luego la de la tinción negativa. Pero el fibrinógeno no se tiñe bien por este último método; ello explica que en los 10 años subsiguientes se considerara simple artefacto la estructura trinodular observada por Hall y Slayter. El fibrinógeno, vino a decirse, semeja un balón de playa hinchado que se abría durante la polimerización para formar fibrina. Mas, recientemente, se ha vuelto a las técnicas de sombreado y, con ellas, a la manifestación de la estructura trinodular del fibrinógeno. Laboratorios ha habido incluso en donde se ha llegado a visualizar la estructura trinodular mediante la tinción negativa.

Muchos otros experimentos demuestran la estructura trinodular (o, al menos, de múltiples dominios) del fibrinógeno, particularmente cuando se consideran en su conjunto. Cabe citar la difracción de rayos X, distintas determinaciones hidrodinámicas (tales como la ultracentrifugación analítica) y otras técnicas fisicoquímicas, así como investigaciones en las que se fragmenta enzimáticamente la molécula de fibrinógeno. En los últimos años, se ha desvelado la estructura primaria de la molécula de fibrinógeno humano (la secuencia completa de aminoácidos y la forma de interconectarse las cadenas). Ello confirma, en detalle, un tipo particular de estructura de múltiples dominios.

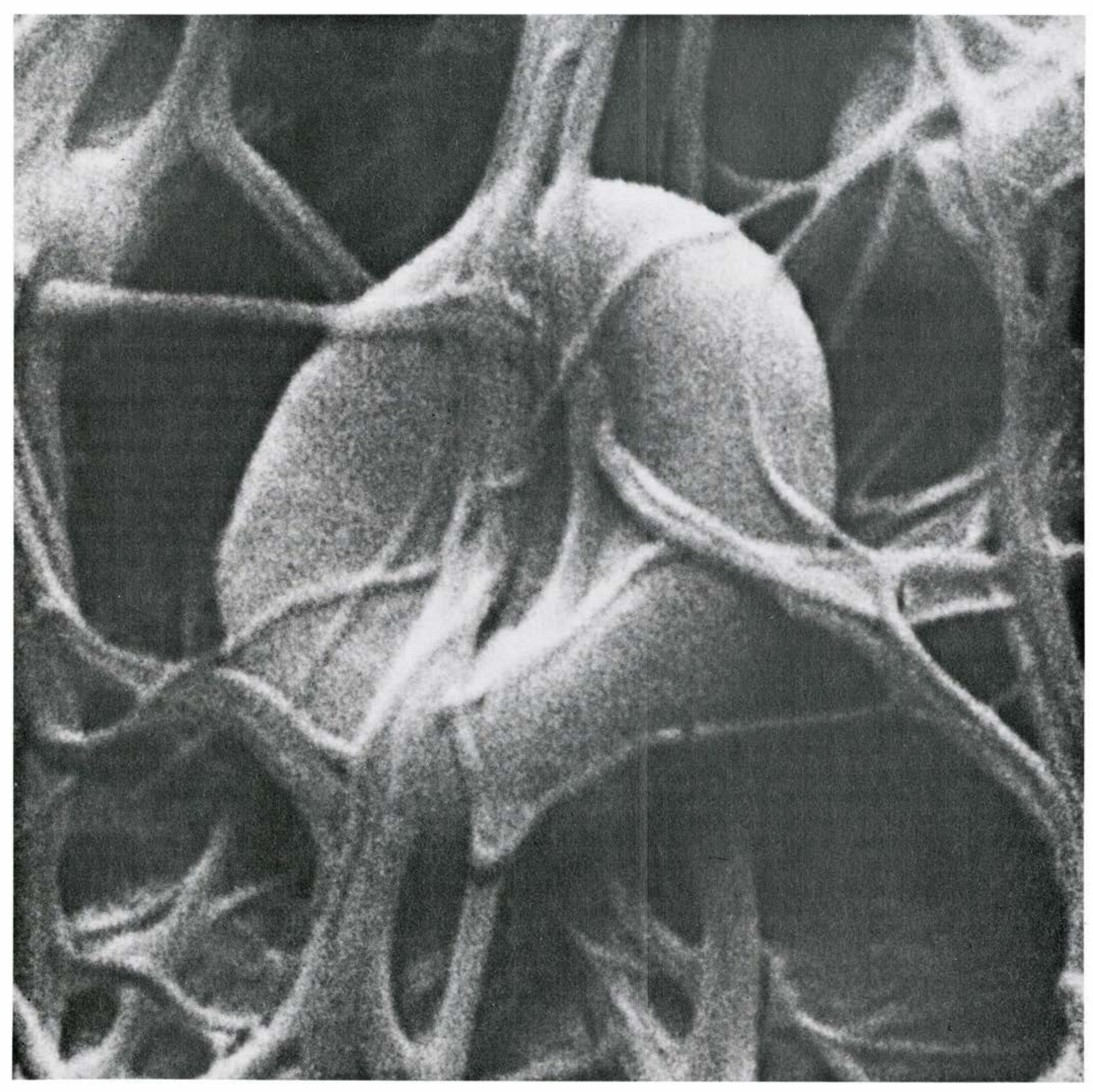
Hacia 1930, W. T. Astbury, de la Universidad de Leeds, inició una serie de estudios exhaustivos sobre la estructura de las proteínas fibrosas por el método de difracción de rayos X. En un material fibroso, las moléculas suelen alinearse en sentido longitudinal; se repite, pues, la distancia de un aminoácido a otro de forma regular. La difracción de rayos X nos revela esa estructura cuando el haz incidente es perpendicular al eje de la fibra. En los primeros años de la década de los 40, Kenneth Bailey colaboraba con Astbury en el establecimiento de las diferencias entre el fibrinógeno y la fibrina. Sus investigaciones demostraron que las dos proteí-

nas presentaban idénticos patrones de difracción que, a su vez, eran indistinguibles de los observados antes por Astbury en ciertas proteínas fibrosas, como la queratina y la miosina.

A raíz de estos experimentos y de otras investigaciones, el fibrinógeno y la fibrina se incluyeron en un grupo de proteínas caracterizado por la repetición de una determinada distancia. Ese intervalo resultó característico de las "espirales enrolladas", donde dos o más cadenas proteicas, arrolladas o superenrolladas entre sí, forman una hélice alfa. Las micrografías electrónicas de Hall y Slayter mostraban una molécula compuesta por tres glóbulos alineados, aunque no se distinguían las conexiones entre ellos, lo que indicaba que su espesor era inferior a los 15 angstrom. Carolyn Cohen, actualmente en la Universidad de Brandeis, sugirió que las conexiones quizá fueran espirales enrolladas.

En las décadas siguientes (de 1940 a 1960), el fibrinógeno, al igual que muchas otras proteínas, se sometió a múltiples estudios fisicoquímicos. Se determinó con precisión su peso molecular, 340.000 dalton, lo que corresponde a unos 3000 aminoácidos. Se investigó asimismo la forma de la molécula. Pero aquí los datos fisicoquímicos valían tanto para una molécula alargada de estructura asimétrica como para una estructura hidratada en alto grado y simétrica. De lo que se desprendía que los datos concordaban también con una molécula compuesta de nódulos enlazados.

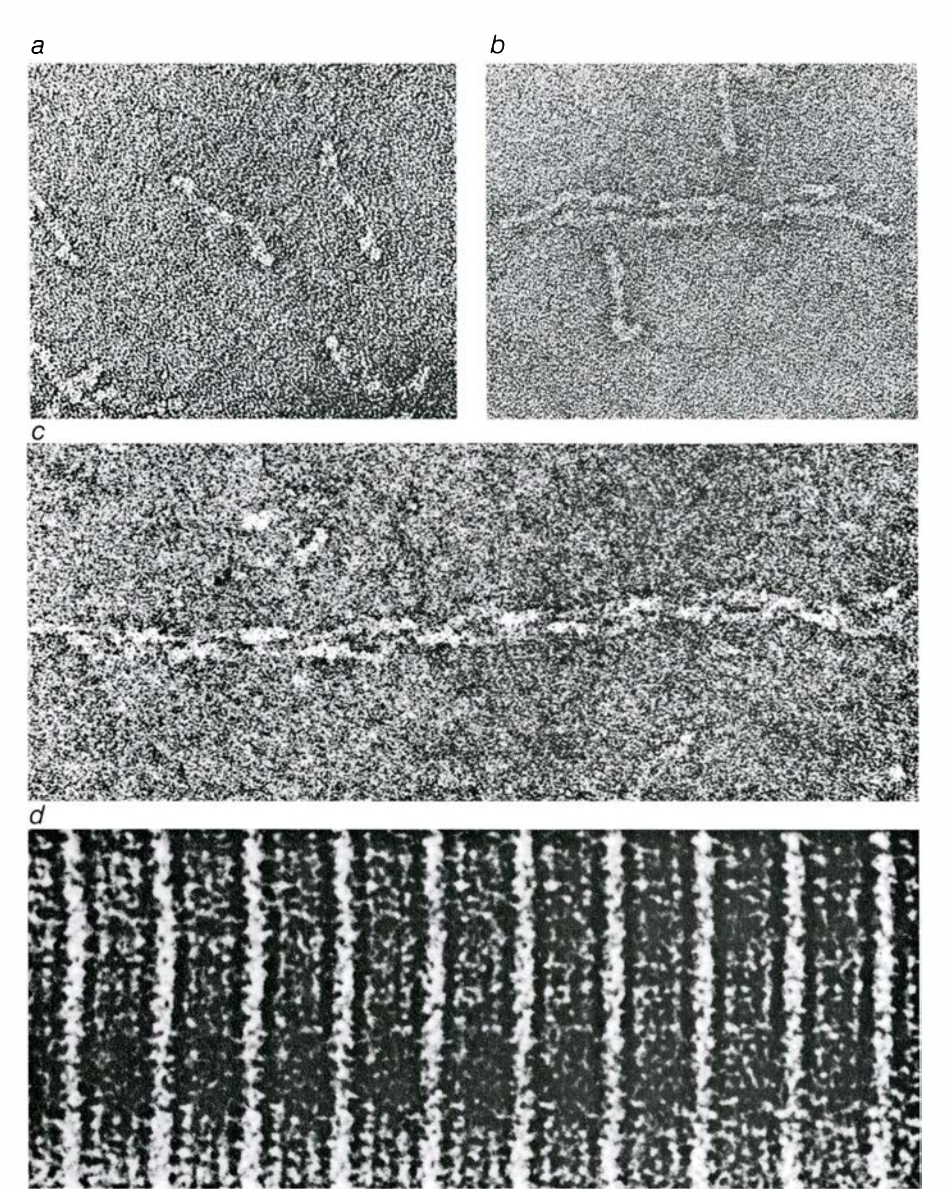
Se recabó información adicional de las hidrólisis enzimáticas. Ya en 1961, investigadores del Instituto Pasteur de París decidieron fragmentar la



HEBRAS DE FIBRINA, ampliadas unas 20.000 veces, atrapando un hematíe, o glóbulo rojo sanguíneo. Cada hebra es un agregado de monómeros de fibrina, formado por la acción del enzima trombina sobre moléculas de fibrinógeno, que polimerizan. El coágulo se formó en un catéter intravenoso implantado cerca del corazón del paciente. La micrografía electrónica es de Emil O. Bernstein y Eila Kairinen, del Instituto de Investigación Gillette.

molécula de fibrinógeno, de gran tamaño, en componentes manejables, utilizando para ello la digestión con plasmina. Obtuvieron así una serie de fragmentos centrales. Los separaron, en base a su carga, por cromatografía de intercambio de iones, y a los fragmentos principales les asignaron, respectivamente, una letra, de la *A* a la *E*. Los fragmentos *D* y *E* constituían el grueso de la masa recuperada; del fragmento *D* había una cantidad doble que del fragmento *E*.

Tales fragmentos, y otra serie producida por digestión con el enzima tripsina, fueron investigados exhaustivamente por Elemer Mihalyi y colaboradores, del Instituto Nacional del Corazón, Pulmones y Sangre. Llegaron a dos conclusiones fundamentales. Primera: la arquitectura del fibrinógeno le hace susceptible a un tipo de hidrólisis que es independiente del enzima atacante. Es la vulnerabilidad del sustrato, más que la especificidad del enzima, la responsable de la hidrólisis. Lo que cabía esperar de una estructura constituida por glóbulos enlazados. Segunda conclusión: los glóbulos tienen una cierta independencia en cuanto a su rotación en la molécula intacta; es decir, en disolución su orientación relativa no es rígida. A estos experimentos siguieron las investigaciones llevadas a cabo por Victor J. Marder, del Instituto Nacional de Artritis y Enfermedades Metabólicas, que se ocupó de la cinética de la hidrólisis del fibrinógeno por la plasmina. La caracterización de los productos intermediarios de la hidrólisis indicaba que el enzima cortaba sucesivamente los enlaces existentes entre los dominios de una molécula integrada por tres porciones globulares.



FIBRINOGENO Y FIBRINA, observados por tinción negativa con acetato de uranilo. Las micrografías electrónicas, realizadas por Robley C. Williams, de la Universidad de California en Berkeley, muestran imágenes ampliadas 375.000 (*a, b, c*) y 500.000 (*d*). Las moléculas de fibrinógeno, tratadas con iones calcio, evidencian una estructura triglobular (*a*). La trombina transforma el fibrinógeno en monómeros de fibrina. En un estadio inicial de la formación del coágulo (*b*) se observan dos cadenas de tres monómeros cada una. Debido al solapamiento de las dos cadenas hay entre ellas un desfase de medio monómero. En un estadio ulterior (*c*) se ha formado ya una hebra de dos cadenas; las regiones claras que se aprecian a lo largo de la hebra representan la unión, extremo con extremo, de los dominios terminales de dos monómeros contiguos. En un segmento de una hebra gruesa de fibrina, totalmente desarrollada (*d*), los monómeros de fibrina se orientan horizontalmente. Las estriaciones verticales prominentes, separadas unos 230 angstrom (la mitad de la longitud de un monómero) se deben a los dominios terminales, que coinciden a lo largo de la estructura tridimensional de la hebra. La fotografía se realizó exponiendo un negativo diez veces, desplazándolo entre las exposiciones un espacio equivalente a la distancia entre estriaciones. La técnica reduce el "ruido" de la micrografía y conserva, al tiempo que confirma, la estructura periódica.

Antes de volver a los estudios de la secuencia de aminoácidos conviene hacer referencia a otros dos datos estructurales. El análisis del punto de fusión de las proteínas corrobora la estructura de múltiples dominios. Si una proteína se calienta en condiciones cuidadosamente controladas, se alcanza una temperatura a la cual la proteína "se funde", es decir, se desnaturaliza. Las cadenas de aminoácidos, que en principio estaban dispuestas en una estructura tridimensional muy precisa, se colapsan. John W. Donovan y Ronald A. Beardslee, del Centro de Investigación de la Región Occidental del Departamento de Agricultura de los Estados Unidos, observó que dos proteínas en íntimo contacto se funden a una determinada temperatura, que es diferente del punto de fusión de cada una de las dos proteínas por separado. Donovan demostró, con Mihalyi, que el fibrinógeno era único a tal respecto. Mostraba dos puntos de fusión: uno para los dos fragmentos *D* y otro para el fragmento *E*, con lo que puso de manifiesto que ambos dominios no se encontraban en íntimo contacto.

Por último, Nancy M. Tooney y Carolyn Cohen estudiaron la estructura de microcristales preparados a partir de fibrinógeno parcialmente digerido, analizándolos por difracción óptica y microscopía electrónica. Sus resultados demostraron que la unidad fundamental del fibrinógeno debía medir unos 450 angstrom, lo que concordaba con los 475 angstrom que determinaran Hall y Slayter por microscopía electrónica y también con muchos de los datos hidrodinámicos. Recientemente, Cohen ha obtenido cristales de fibrinógenos modificados. Si bien los resultados son provisionales, parecen concordar con una molécula que ocuparía un "cilindro de influencia" de 450 angstrom de longitud por 90 de diámetro.

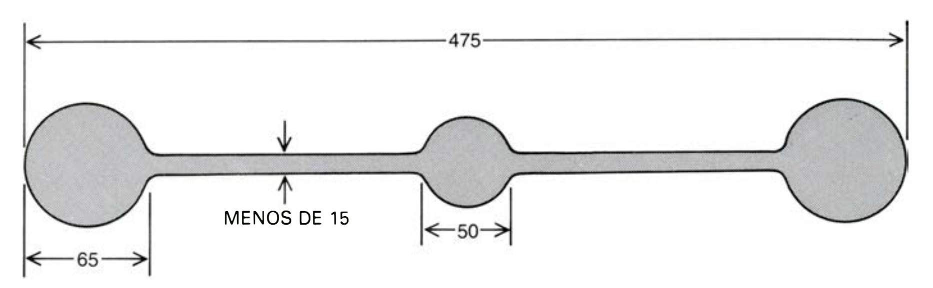
Establecer la secuencia de aminoácidos de una molécula del tamaño del fibrinógeno constituye una tarea ímproba. Pero muchos aspectos de la química del fibrinógeno sólo admitían interpretación a través de la secuencia de aminoácidos; y así, varios laboratorios se aprestaron a determinarla.

Las cadenas proteicas poseen un extremo con un grupo amino (NH_2) y otro con un grupo carboxilo (COOH). Al determinar los aminoácidos aminoterminales de las cadenas de fibrinógeno, se identificaron tres distintos, existiendo dos copias de cada uno de ellos en la molécula intacta de 340.000 dalton. Se aislaron entonces tres cadenas diferentes. El peso molecular de las mismas, en conjunto, representaba la mitad del peso molecular de la molécula intacta. Estos hechos pusieron claramente de manifiesto que la molécula constaba de dos grupos de tres cadenas distintas conectadas por puentes disulfuro. (Uno de los 20 aminoácidos, la cisteína, puede unirse a otra cisteína mediante un puente disulfuro.) Las cadenas se denominaron alfa, beta y gamma.

Los primeros estudios sobre la secuencia del fibrinógeno humano los realizó Birger Blombäck, en el Instituto Karolinska de Estocolmo, en los últimos años de la década de los 60. Blombäck y sus colaboradores no trataron el fibrinógeno con proteasas sino con bromuro de cianógeno, que rompe las proteínas únicamente por determinados enlaces. Aislaron un fragmento que contenía todos los aminoácidos aminoterminales de las tres cadenas distintas. El fragmento era un dímero (un compuesto formado por dos unidades trabadas por puentes disulfuro); además, todos los grupos aminoterminales de la molécula de fibrinógeno formaban un bloque en la molécula intacta.

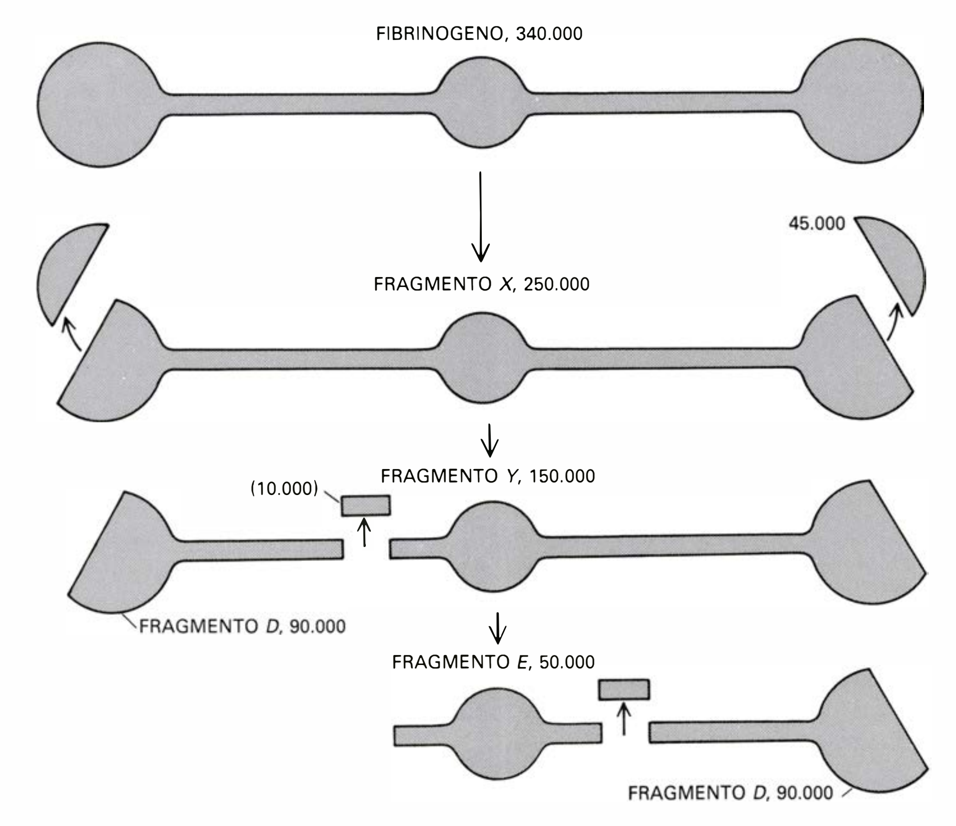
El fragmento representaba sólo alrededor del 15 por ciento de la masa del fibrinógeno; pero contenía casi la mitad de las unidades de dos cisteínas (cistinas) que forman puentes disulfuro, por lo que se le denominó lazo disulfuro. La secuencia del fragmento, que portaba 247 aminoácidos en cada mitad, fue determinada por el grupo del Instituto Karolinska. Experimentos posteriores, realizados por Marder, mostraron que el lazo disulfuro era muy similar al fragmento *E*. Ambos son dímeros y muestran reacciones cruzadas en los ensayos inmunológicos. Hay dos fragmentos *D* por cada fragmento *E*. El fragmento *E* (o lazo disulfuro) es el dominio central del fibrinógeno; los dos fragmentos *D*, los dominios terminales.

Tras ese trabajo inicial de Blombäck, la investigación se detuvo por un tiempo. En su reanudación participaron dos nuevos laboratorios: el de Agnes Henschen, en el Instituto Max Plank de Bioquímica en Munich, y el nuestro, en la Universidad de California en San Diego. Después de varios años de intenso esfuerzo se estableció la estructura covalente global de la molécula. En su mayor parte los resultados se obtuvieron independientemente en ambos laboratorios. La cadena alfa consta de 610 aminoácidos, de 461 la beta y 411 la gamma, de un total de 1482 aminoácidos que componen cada mitad del dímero. Al peso molecular, calculado en 329.842, hay que añadir 10.000 dalton más, correspondientes a dos residuos carbohidratados en cada monómero.



MODELO TRINODULAR, propuesto en 1959 por Cecil E. Hall y Henry S. Slayter, del Instituto de Tecnología de Massachusetts, para la molécula de fibrinógeno. Las dimensiones se dan en angstrom. Este primer modelo, luego confirmado, se basaba, fundamentalmente, en el aspecto que ofrecía la molécula al observarla, al microscopio electrónico, utilizando la técnica de sombreado con metales pesados.

Consideraremos ahora algunos de los hechos cruciales revelados por la secuencia, en particular los que tienen que ver con la estructura tridimensional. La idea inicial de que las tres cadenas podían ser homólogas (lo suficientemente semejantes para derivar de una cadena ancestral común) se ha ido confirmando a medida que se acumulaban datos sobre su secuencia. Al igual que en otras proteínas, la homología se deduce especialmente de la localización de las cisteínas. Todas ellas, como demostró Henschen, participan en la formación de puentes disulfuro.



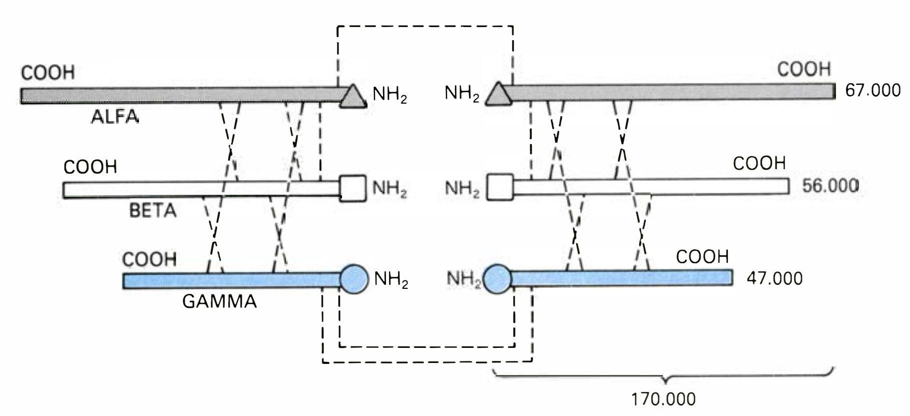
DIGESTION DEL FIBRINOGENO con el enzima plasmina, estudiada por Victor J. Marder en 1968. El enzima daba lugar, sucesivamente, a los productos intermediarios *X* e *Y* y a los fragmentos nucleares *D* y *E*, mostrados aquí junto a su peso molecular. Los fragmentos y el curso temporal de su formación coinciden con la hidrólisis enzimática de las uniones que enlazan los dominios de una molécula triglobular.

El aspecto más interesante de los puentes disulfuro, a mi entender, se hizo patente durante la secuenciación de la cadena alfa. Blombäck observó que las tres cadenas distintas poseían una serie poco común de cisteínas en la siguiente secuencia: cisteína-*X*-*Y*-*Z*-cisteína. Observamos una segunda secuencia de este tipo en la cadena alfa, 111 aminoácidos más allá. ¿Poseían también las otras dos cadenas secuencias similares en las mismas posiciones relativas? En pocos días obtuvimos una respuesta afirmativa; costó mucho más dar con una demostración. Las tres cadenas poseen dos cisteínas en una secuencia de cinco aminoácidos separadas por 111 aminoácidos (en las cadenas alfa y gamma) o por 112 (en la cadena beta). A partir de estos datos se plantearon dos cuestiones estructurales básicas: ¿cómo se encuentran dispuestos los puentes disulfuro entre estas cisteínas particulares y cuál es la estructura tridimensional de los 111 o 112 aminoácidos que separan las dos secuencias que contienen cisteína en cada cadena?

Antes incluso de empezar nuestros experimentos teníamos una respuesta al primer problema, basándonos en consideraciones evolutivas. Creíamos que en una molécula de fibrinógeno ancestral existirían tres cadenas idénticas, que habrían variado con el tiempo por mutación de sus genes. Resulta evidente que la misma secuencia en el mismo ambiente debe tener la misma estructura tridimensional, por lo que el conjunto de la molécula debería poseer una triple simetría. Ello nos hizo suponer que el primer puente disulfuro uniría la primera cisteína de una cadena con la segunda cisteína de la segunda cadena, la primera cisteína de la segunda cadena con la segunda cisteína de la tercera cadena y, finalmente, la primera cisteína de la tercera cadena con la segunda cisteína de la primera cadena, cerrando así un "anillo disulfuro". Cuando construimos un modelo de tal estructura, las cadenas de cisteína encajaban en su sitio y los puentes disulfuro, en efecto, formaban un anillo. Todos los datos bioquímicos acumulados desde entonces apoyan el modelo propuesto.

La naturaleza de las secuencias que separaban los anillos disulfuro encerraba un interés mayor. La observación de las secuencias reveló que los aminoácidos con cadenas apolares (hidrofóbicas) se repetían periódicamente a lo largo de la secuencia, lo que sugería una disposición helicoidal. Los estudios realizados, por medio de ordenadores, para investigar la probabilidad de que una secuencia particular dé lugar a una determinada conformación, pusieron de manifiesto que la mayoría del segmento existente entre los anillos disulfuro debía de adoptar la conformación de una hélice alfa. Pareció claro que tales segmentos habían de corresponder a las espirales enrolladas. Los modelos nos permitieron mostrar que las secuencias conducían a la formación de una fibra de tres hebras, en la que todos los grupos laterales de naturaleza polar (hidrofílicos) se dirigían hacia el medio acuoso y la mayoría de las cadenas laterales hidrofóbicas se orientaban hacia el interior para constituir un núcleo apolar. Por último, se determinó la distancia entre los anillos disulfuro. Mientras que en una hélice alfa convencional cada aminoácido avanza aproximadamente 1,5 angstrom, el superenrollamiento acorta la distancia a 1,43 angstrom. Si multiplicamos este valor por 111, la distancia que media entre los anillos disulfuro resulta ser de 158,7 angstrom. Ese intervalo concuerda con el del modelo de Hall-Slayter, en el que los glóbulos están separados por una distancia de unos 150 angstrom.

Las únicas regiones de importante estructura alfa helicoidal en las cadenas parecen ser las comprendidas entre anillos disulfuro. Las regiones carboxiterminales de las cadenas beta y gamma, que muestran una gran homología, se manifiestan muy plegadas, constituyendo regiones compactas que forman la mayor parte de los fragmentos terminales *D*. En cambio, la región carboxiterminal de las cadenas alfa se encuentra desplegada, tiene naturaleza polar y constituye una protuberancia que hidrolizan fácilmente distintas proteasas. El fibrinógeno carente de estas regiones de cadena alfa guarda un parecido estrecho con el producto intermediario que Marder denominó fragmento *X*.



ESTRUCTURA DE SEIS CADENAS del fibrinógeno, establecida a comienzos de la década de 1970. Se observó que los grupos amino (NH_2) terminales estaban enlazados a tres aminoácidos distintos, y que había dos copias de cada "grupo terminal". La ruptura de los puentes disulfuro (*líneas a trazos*) permitió aislar tres cadenas diferentes: alfa, beta y gamma, cuyo peso molecular total, de 170.000, representaba la mitad del peso de la molécula. Había, por tanto, dos copias de cada cadena. Se observaron también numerosos puentes disulfuro entre los extremos aminoterminales, coincidiendo con el dominio central.

A partir de los datos estructurales se planteará ahora la cuestión de cómo se relacionan estructura y función, tanto en lo que respecta a la polimerización como a la fibrinolisis. Hace 30 años se descubrió que el fenómeno desencadenante de la transformación del fibrinógeno en fibrina era la liberación de dos pares de pequeños péptidos polares (cortas cadenas proteicas) catalizada por la trombina. Se trata de los fibrinopéptidos *A* y *B*, que se encuentran en los extremos aminoterminales de las cadenas alfa y beta, respectivamente. Las moléculas de fibrinógeno se convierten así en monómeros de fibrina, que se unen espontáneamente para formar una red de fibrina. ¿Cuál es la razón para que polimericen los monómeros de fibrina y no lo hagan las moléculas de fibrinógeno? La interpretación inicial argumentaba que los fibrinopéptidos, portadores de una notable carga negativa, evitaban la interacción de las moléculas de fibrinógeno por repulsión electrostática. En 1952, John D. Ferry, de la Universidad de Wisconsin en Madison, propuso que la liberación de los fibrinopéptidos cargados negativamente permitiría la aproximación entre monómeros, estableciéndose diferentes interacciones físicas débiles de corto alcance.

Ferry hizo otra notable observación sobre la polimerización. Se sabía que, en las micrografías electrónicas, los haces de fibrina formaban un característico patrón en bandas. La distancia repetitiva era de unos 230 angstrom. Puesto que ese intervalo espacial constituía la mitad aproximada de un monómero de fibrina, Ferry sugirió que la polimerización debía tener lugar por apareamien-

FIBRINOPEPTIDO *A*

1	Ala	Asp	Ser	Gly	Glu	Gly	Asp	Phe	Leu	Ala	Glu	Gly	Gly	Gly	Val	Arg	Gly	Pro	Arg	Val	Val	Glu	Arg	His	Gln
26	Ser	Ala	Cys	Lys	Asp	Ser	Asp	Trp	Pro	Phe	Cys	Ser	Asp	Glu	Asp	Trp	Asn	Tyr	Lys	Cys	Pro	Ser	Gly	Cys	Arg
51	Met	Lys	Gly	Leu	Ile	Asp	Glu	Val	Asn	Gln	Asp	Phe	Thr	Asn	Arg	Ile	Asn	Lys	Leu	Lys	Asn	Ser	Leu	Phe	Glu
76	Tyr	Gln	Lys	Asn	Asn	Lys	Asp	Ser	His	Ser	Leu	Thr	Thr	Asn	Ile	Met	Glu	Ile	Leu	Arg	Gly	Asp	Phe	Ser	Ser
101	Ala	Asn	Asn	Arg	Asp	Asn	Thr	Tyr	Asn	Arg	Val	Ser	Glu	Asp	Leu	Arg	Ser	Arg	Ile	Glu	Val	Leu	Lys	Arg	Lys
126	Val	Ile	Gln	Lys	Val	Gln	His	Ile	Gln	Leu	Leu	Gln	Lys	Asn	Val	Arg	Ala	Gln	Leu	Val	Asp	Met	Lys	Arg	Leu
151	Glu	Val	Asp	Ile	Asp	Ile	Lys	Ile	Arg	Ser	Cys	Arg	Gly	Ser	Cys	Ser	Arg	Ala	Leu	Ala	Arg	Glu	Val	Asp	Leu
176	Lys	Asn	Tyr	Glu	Asp	Gln	Gln	Lys	Gln	Leu	Glu	Gln	Val	Ile	Ala	Lys	Asp	Leu	Leu	Pro	Ser	Arg	Asp	Arg	Gln
201	His	Leu	Pro	Leu	Ile	Lys	Met	Lys	Pro	Val	Pro	Asn	Leu	Val	Pro	Gly	Asn	Phe	Lys	Ser	Gln	Leu	Gln	Lys	Val
226	Pro	Pro	Glu	Trp	Lys	Ala	Leu	Thr	Asp	Met	Pro	Gln	Met	Arg	Met	Glu	Leu	Glu	Arg	Pro	Gly	Gly	Asn	Glu	Ile
251	Thr	Arg	Gly	Gly	Ser	Thr	Ser	Tyr	Gly	Thr	Gly	Ser	Glu	Thr	Glu	Ser	Pro	Arg	Asn	Pro	Ser	Ser	Ala	Gly	Ser
276	Trp	Asn	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Thr	Gly	Asn	Arg	Asn	Pro	Gly	Ser	Ser	Gly	Thr	Gly	Ser	Gly	Ala
301	Thr	Trp	Lys	Pro	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Thr	Gly	Ser	Trp	Asn	Ser	Gly	Ser	Ser	Gly	Thr	Gly	Ser	Thr
326	Gly	Asn	Gln	Asn	Pro	Gly	Ser	Pro	Arg	Pro	Gly	Ser	Thr	Gly	Thr	Trp	Asn	Pro	Gly	Ser	Ser	Glu	Arg	Gly	Ser
351	Ala	Gly	His	Trp	Thr	Ser	Glu	Ser	Ser	Val	Ser	Gly	Ser	Thr	Gly	Gln	Trp	His	Ser	Glu	Ser	Gly	Ser	Phe	Arg
376	Pro	Asp	Ser	Pro	Gly	Ser	Gly	Asn	Ala	Arg	Pro	Asn	Asp	Pro	Asn	Trp	Gly	Thr	Phe	Glu	Glu	Val	Ser	Gly	Asn
401	Val	Ser	Pro	Gly	Thr	Arg	Arg	Glu	Tyr	His	Thr	Glu	Lys	Leu	Val	Thr	Ser	Lys	Gly	Asp	Lys	Glu	Leu	Arg	Thr
426	Gly	Lys	Glu	Lys	Val	Thr	Ser	Gly	Ser	Thr	Thr	Thr	Thr	Arg	Arg	Ser	Cys	Ser	Lys	Thr	Val	Thr	Lys	Thr	Val
451	Ile	Gly	Pro	Asp	Gly	His	Lys	Glu	Val	Thr	Lys	Glu	Val	Val	Thr	Ser	Glu	Asp	Gly	Ser	Asp	Cys	Pro	Glu	Ala
476	Met	Asp	Leu	Gly	Thr	Leu	Ser	Gly	Ile	Gly	Thr	Leu	Asp	Gly	Phe	Arg	His	Arg	His	Pro	Asp	Glu	Ala	Ala	Phe
501	Phe	Asp	Thr	Ala	Ser	Thr	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Ser	Pro	Met	Leu	Gly	Glu	Phe	Val	Ser	Glu	Thr
526	Glu	Ser	Arg	Gly	Ser	Glu	Ser	Gly	Ile	Phe	Thr	Asn	Thr	Lys	Glu	Ser	Ser	Ser	His	His	Pro	Gly	Ile	Ala	Glu
551	Phe	Pro	Ser	Arg	Gly	Lys	Ser	Ser	Ser	Tyr	Ser	Lys	Gln	Phe	Thr	Ser	Ser	Thr	Ser	Tyr	Asn	Arg	Gly	Asp	Ser
576	Thr	Phe	Glu	Ser	Lys	Ser	Tyr	Lys	Met	Ala	Asp	Glu	Ala	Gly	Ser	Glu	Ala	Asp	His	Glu	Gly	Thr	His	Ser	Thr
601	Lys	Arg	Gly	His	Ala	Lys	Ser	Arg	Pro	Val															

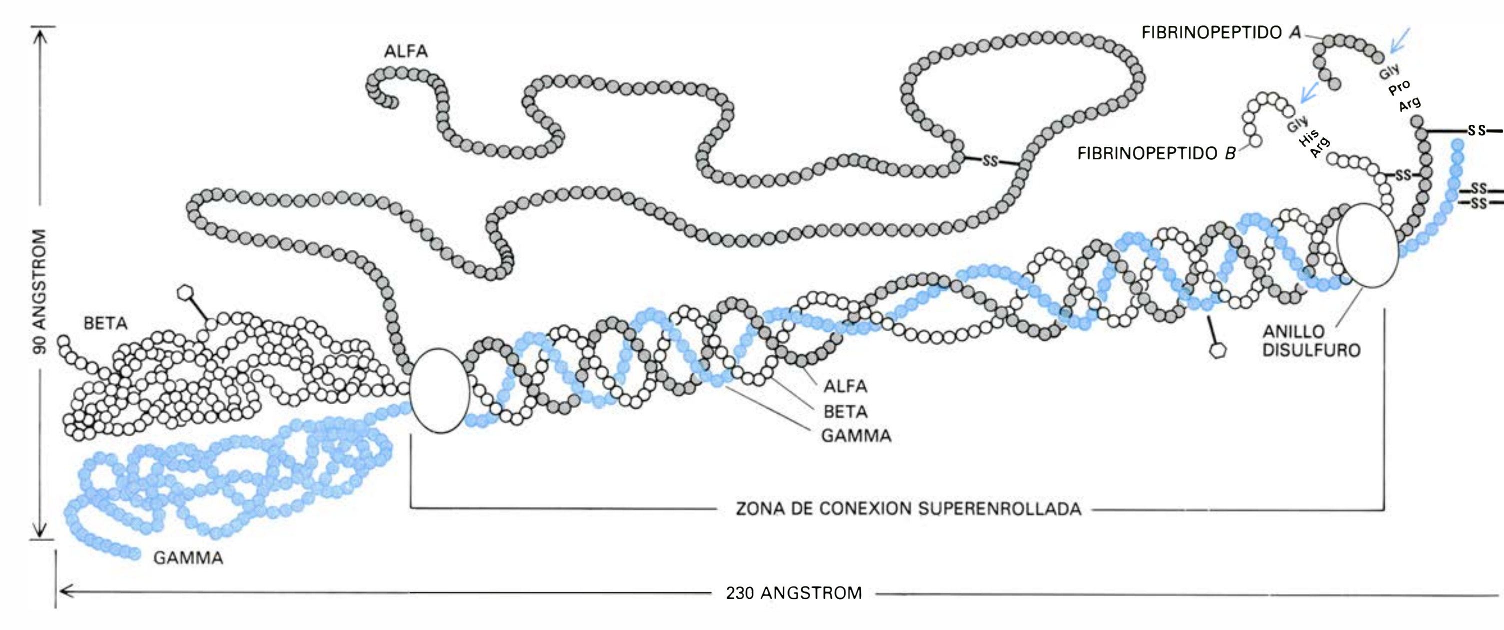
Ala	Alanina	**Leu**	Leucina
Arg	Arginina	**Lys**	Lisina
Asn	Asparagina	**Met**	Metionina
Asp	Aspartato	**Phe**	Fenilalanina
Cys	Cisteína	**Pro**	Prolina
Gln	Glutamina	**Ser**	Serina
Glu	Glutamato	**Thr**	Treonina
Gly	Glicina	**Trp**	Triptófano
His	Histidina	**Tyr**	Tirosina
Ile	Isoleucina	**Val**	Valina

SECUENCIA COMPLETA DE AMINOACIDOS de la cadena alfa, determinada en su mayor parte en el laboratorio del autor. Las abreviaturas correspondientes a los 20 aminoácidos se indican a la izquierda. La flecha coloreada muestra el punto donde la trombina hidroliza la cadena, liberando el fibrinopéptido *A* del extremo aminoterminal. La zona sombreada más oscura indica la zona de conexión constituida por "hélices enrolladas". Las ocho cisteínas que forman parte de los puentes disulfuro aparecen en color.

to lateral con "un solapamiento parcial que da lugar a dos cadenas paralelas con uniones escalonadas". Esta interpretación concuerda con los datos obtenidos en las investigaciones del polímero en crecimiento llevadas a cabo mediante dispersión de la luz.

Bailey y Astbury no encontraron diferencias entre los patrones de difracción de rayos X del fibrinógeno y de la fibrina, lo que les llevó a la conclusión de que "la fibrina no es más que una forma insoluble del fibrinógeno, sin ningún cambio fundamental en la estructura molecular". Recientes estudios calorimétricos han permitido concluir a Donovan y Mihalyi que no existían diferencias significativas en los puntos de fusión del fibrinógeno y la fibrina. A la vez que se confirmaba la semejanza, en lo esencial, del fibrinógeno y el monómero de fibrina, se demostraba así que las interacciones entre monómeros de fibrina no eran particularmente "cooperativas". Debían de ser escasas y bastante rígidas. Aparentemente, dos monómeros de fibrina estarían unidos entre sí de forma más parecida a los danzantes de sardana, que bailan unidos por sus manos, que a quienes bailan "arrimaditos" un tango, mejilla contra mejilla.

Pasemos de nuevo revista a algunas de las hipótesis iniciales sobre la polimerización, a la luz de las secuencias de aminoácidos, empezando por las interacciones electrostáticas. La molécula de fibrinógeno tiene, en el hombre, un exceso de 26 cargas negativas. Ese sobrante de cargas no se encuentra repartido uniformemente a lo largo de la molécula. La mayor parte se encuentra en el dominio central. La liberación de los dos fibrinopéptidos *A* reduce el exceso de cargas negativas del dominio central de menos 8 a menos 1. Cuando se liberan los fibrinopéptidos *B*, el dominio central queda polarizado con un exceso de cargas positivas de más 5. Las regiones constituidas por hélices enrolladas tienen aproximadamente el mismo número de cadenas laterales cargadas positivamente que negativamente, lo que está de acuerdo con la idea de que tales zonas de conexión no están implicadas en las interacciones intermoleculares, pero permanecen, sin embargo, accesi-



MODELO DETALLADO DEL FIBRINOGENO, diseñado por el autor y colaboradores a partir de los datos de la secuencia de aminoácidos. La molécula es un dímero. La línea vertical a trazos muestra su eje de simetría. Está compuesta de dos juegos de tres cadenas: alfa (*gris*), beta (*blanca*) y gamma (*coloreada*). Las cadenas están unidas por 29 puentes disulfuro (*S-S*), 13 de ellos en cada mitad del dímero y 3 estableciendo la conexión de las dos mitades; los "anillos disulfuro" tienen tres de esas uniones, como se muestra en detalle en la ilustración inferior de estas dos páginas. La zona de conexión entre anillos

bles a un posible ataque de la plasmina.

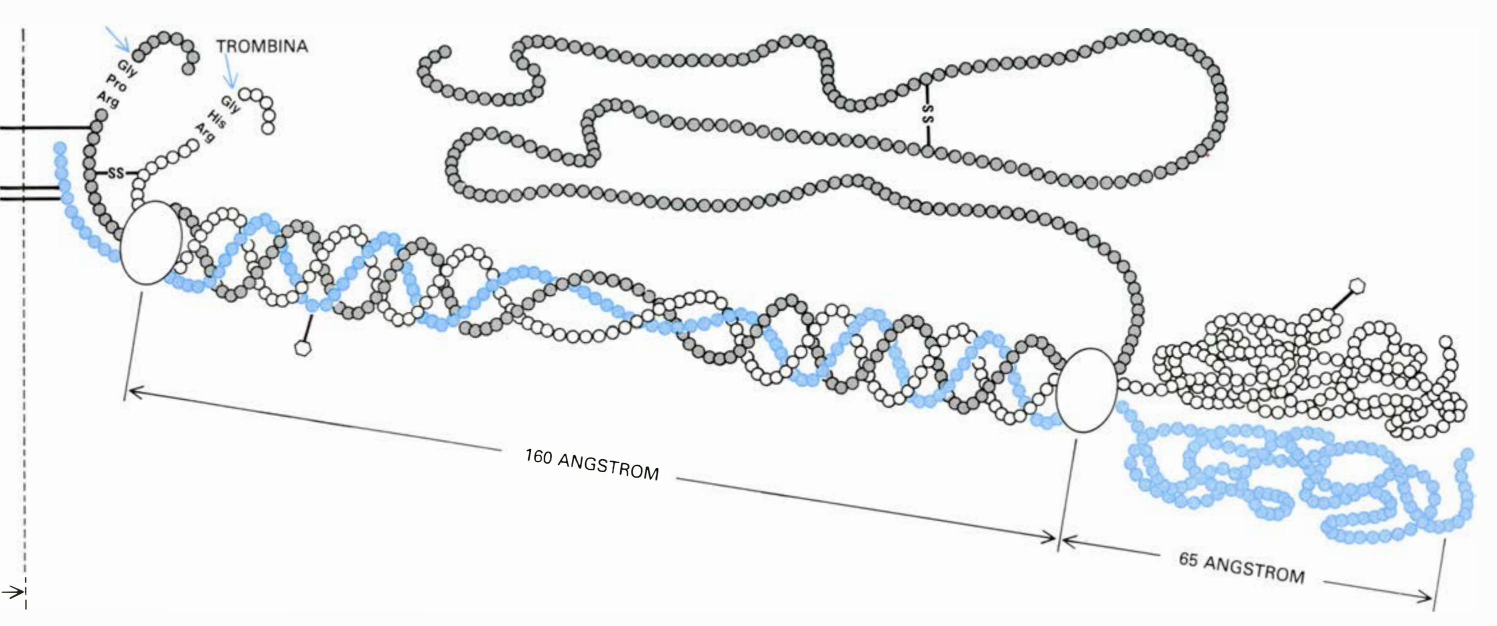
Cada uno de los dominios terminales posee, sin embargo, una carga neta negativa de menos 4. Lo que encaja perfectamente con la posibilidad de que cada dominio terminal establezca una interacción electrostática complementaria con el dominio central, cargado positivamente, dando lugar, durante la polimerización, al solapamiento de los monómeros de fibrina con uniones escalonadas, tal como propuso Ferry a partir del patrón regular en bandas que pudo observar con el microscopio electrónico.

¿Cuál es el mecanismo de la interacción? Una hipótesis simple sería que existen zonas específicas de polimerización en los monómeros, que, en el fibrinógeno, están bloqueadas por los fibrinopéptidos. La idea de zonas específicas de polimerización sugiere la existencia de estructuras complementarias, es decir, sitios dadores y sitios aceptores, o bien "botones" y "ojales". Para un mayor rendimiento y economía bastaría con que los fibrinopéptidos bloquearan a los botones o a los ojales. Supongamos que los ojales están siempre libres. En ese caso, la liberación de los fibrinopéptidos del dominio central puede poner al descubierto una serie de botones que interaccionen con los ojales de los dominios terminales de las moléculas vecinas. Puesto que el monómero inicial tiene dos extremos, puede crecer en cualquiera de los dos sentidos. El polímero en crecimiento tiene un espesor correspondiente a dos moléculas y se acomoda al solapamiento de las moléculas de fibrinógeno.

Partiendo de la idea general de que existe una interacción entre botones y ojales, con los botones situados en el dominio central, bloqueados por fibrinopéptidos, y los ojales siempre expuestos en los dominios terminales, la cuestión que se plantea es cuántos existen y si son todos ellos equivalentes. Hay dos tipos de fibrinopéptidos, *A* y *B*, y se sabe también que se liberan con diferentes velocidades. T. C. Laurent y Blombäck sugirieron hace tiempo que la liberación de los fibrinopéptidos *A*, fenómeno que ocurre en primer lugar en los mamíferos, permite la formación de polímeros constituidos fundamentalmente por interacciones extremo-extremo, mientras que la liberación subsiguiente de fibrinopéptidos *B* permitiría la interacción lateral que conduce a la formación de fibras grue-



SECUENCIA DE AMINOACIDOS de las tres cadenas del segmento que contiene los dos anillos disulfuro (*a*). (Corresponde, desde luego, a una mitad del dímero de fibrinógeno; la otra mitad tiene un segmento equivalente.) En la mayor parte de la región, las cadenas forman una hélice superenrollada, lo que coincide con el patrón de difracción de rayos X obtenido en los estudios iniciales. Las hélices superenrolladas explican también las regiones en forma



disulfuro está constituida, fundamentalmente, por hélices alfa superenrolladas, con una región laxa, que no presenta estructura helicoidal, en el centro de cada zona de conexión. Los dominios terminales de las cadenas beta y gamma están empaquetados de forma compacta. La cadena alfa es hidrofílica y "flota" libremente en el medio acuoso. Dos cadenas carbohidratadas (*hexágonos*) se encuentran unidas a cada mitad. Se indican los aminoácidos situados en los sitios donde la liberación de los fibrinopéptidos forma los "botones" alfa y beta. El dibujo, como los demás del artículo, es de G. V. Kelvin.

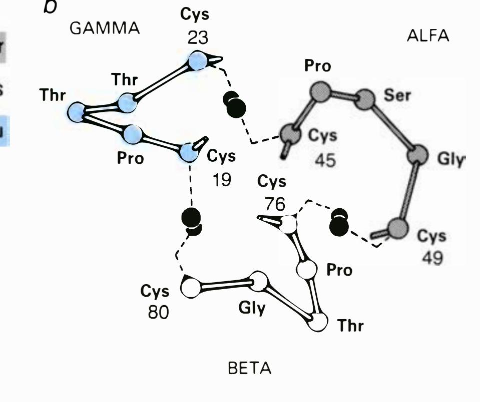
sas. Su idea se basaba en las hipótesis iniciales sobre polimerización y formación de las fibras, propuestas por Ferry y Peter R. Morrison, a la sazón en la Universidad de Harvard. Partiendo de las diferencias de opacidad de los geles de fibrina formados en diferentes condiciones, distinguieron entre coágulos de trama fina, en los que las fibras estaban constituidas por protofibrillas de una o dos moléculas de espesor, y coágulos de fibra gruesa, formados por haces de protofibrillas asociadas lateralmente. De acuerdo con estas ideas debían existir dos tipos diferentes de botones, unos bloqueados por los fibrinopéptidos *A* y otros por los fibrinopéptidos *B*, y también dos tipos de ojales complementarios.

Desde un principio pareció plausible que las regiones aminoterminales expuestas tras la liberación de los fibrinopéptidos por la trombina constituyeran los sitios de polimerización. Decidimos pues, en colaboración con mi discípulo Andrew P. Laudano, sintetizar lo que debían ser los botones: péptidos correspondientes a los segmentos aminoterminales expuestos tras la liberación de los fibrinopéptidos. Observamos que los tripéptidos sintéticos, glicina-prolina-arginina (o péptidos más largos que empezaban por dicha secuencia), correspondientes a la secuencia aminoterminal de la cadena alfa, se unían al fibrinógeno de ciertas preparaciones del fragmento *D*. Tenían también la propiedad de inhibir la polimerización de los monómeros de fibrina. Los botones sintéticos ocupaban los ojales e impedían la interacción normal con los botones del dominio central de los monómeros. Por otra parte, si bien los péptidos correspondientes a la secuencia de la cadena beta (glicina-histidina-arginina) expuestos tras la liberación del fibrinopéptido *B* también se unían al fibrinógeno y al fragmento *D*, no impedían la polimerización de los monómeros de fibrina en el sistema humano. Estudios competitivos pusieron de manifiesto que se hallaban involucrados dos tipos de sitios. Evidentemente, la polimerización inicial dependía exclusivamente de los botones de la cadena alfa.

La localización de los ojales en los dominios terminales se ha determinado recientemente con mayor precisión. Frits Haverkate, de la Universidad de Leiden, demostró que la naturaleza del fragmento *D* generado por la plasmina está influida por los iones calcio. En su

Thr Thr Asn Ile Met Glu Ile Leu Arg Gly Asp Phe Ser Ser Ala Asn Asn Arg Asp Asn Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser
Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys Gln Val Lys Asp Asn Glu Asn Val Val Asn Glu Tyr Ser Ser Glu Leu Glu Lys
Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro Asp Glu Ser Ser Lys Pro Asn Met Ile Asp Ala Ala Thr Leu Lys Ser Arg Lys Met Leu Glu

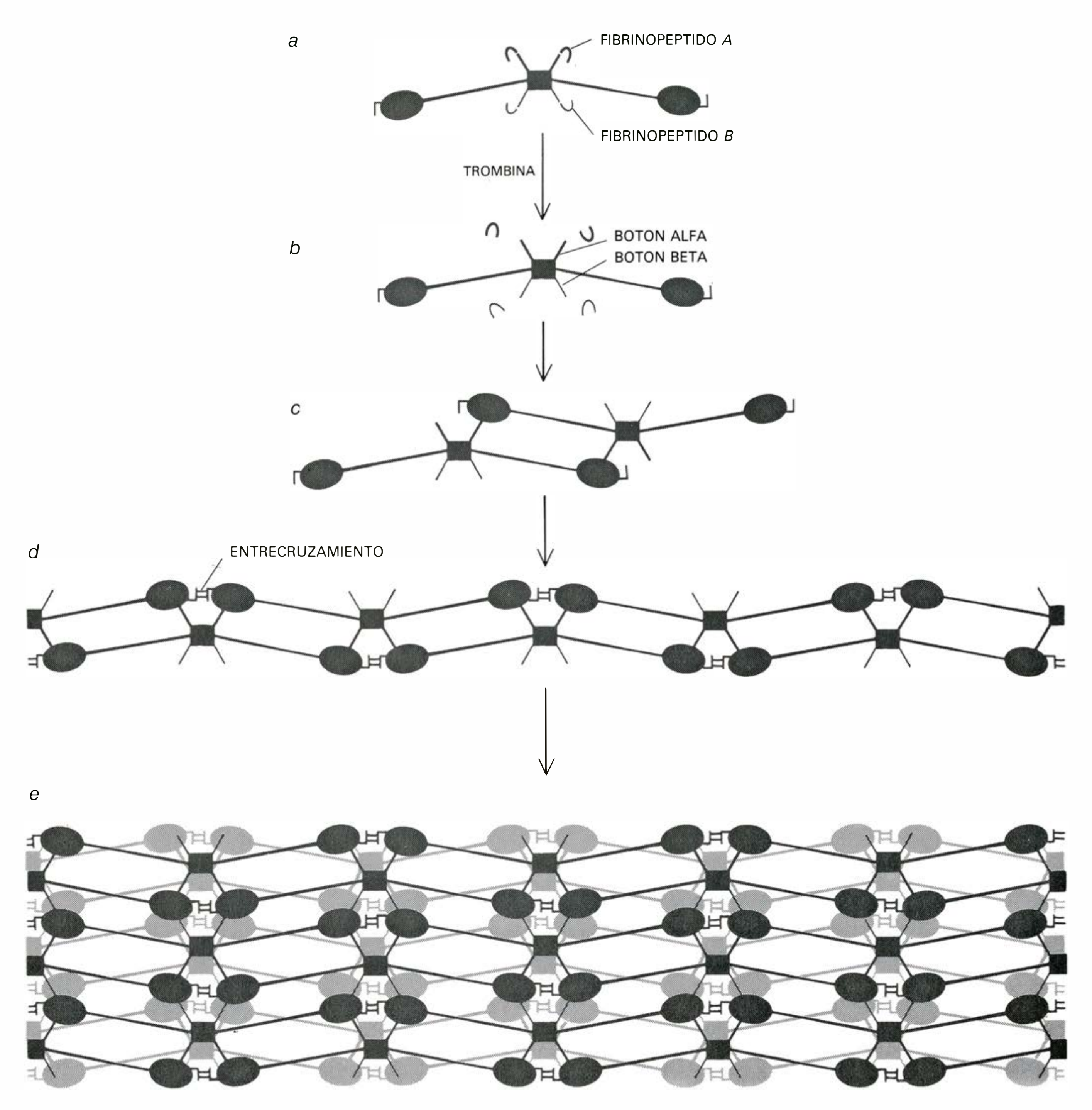
161 165 SEGUNDO ANILLO DISULFURO
Arg Ser Cys Arg Gly Ser Cys ALFA
193 197
Glu Tyr Cys Arg Thr Pro Cys BETA
Ala Gln Cys Gln Glu Pro Cys GAMMA
135 139



de bastón del modelo de Hall-Slayter, es decir, las zonas de conexión entre los dominios globulares central y terminal de la molécula. Las flechas coloreadas indican los puntos de hidrólisis de la plasmina. Los tres enlaces de cada anillo disulfuro se muestran también en color. Un modelo molecular de un anillo disulfuro (*b*), visto desde la región aminoterminal de la molécula, muestra la conformación de las secuencias cisteína-*X*-*Y*-*Z*-cisteína de las tres cadenas.

ausencia, la digestión prosigue mucho más allá, generándose pequeños fragmentos *D* en los que la cadena gamma se encuentra notablemente acortada en su extremo carboxiterminal. Al examinar la unión de los botones sintéticos a los dos tipos citados de fragmento *D*, observamos que los botones alfa se unían solamente a los fragmentos *D* de mayor tamaño, mientras que la unión se abolía al acortarse la cadena gamma. Sin embargo, los botones beta se unían por igual a los dos tipos de fragmento *D*. Evidentemente, el segmento carboxiterminal de la cadena gamma contribuye a formar los ojales para los botones alfa. Recientemente, Stephanie Olexa y Andrei Z. Budzynski, del Hospital Clínico de la Universidad de Temple, aislaron un péptido del extremo carboxiterminal de la cadena gamma que se une al dominio central y que parece ser una parte esencial del ojal del botón alfa.

En la interacción botón-ojal establecida bajo la única influencia de la trombina se hallan implicadas fuerzas de atracción (incluidas interacciones electrostáticas y quizá también enlaces de hidrógeno) que pueden neutralizarse



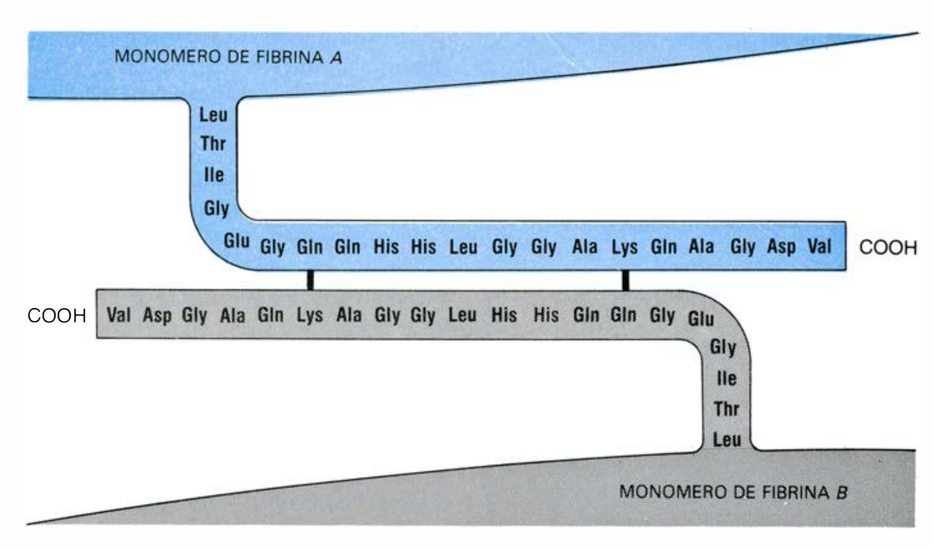
ESQUEMA DE POLIMERIZACION, propuesto de acuerdo con los datos obtenidos por microscopía electrónica y otras técnicas de investigación de la estructura. La molécula de fibrinógeno (*a*) se convierte en el monómero de fibrina (*b*) mediante la liberación de los fibrinopéptidos *A* y *B* a partir de los extremos aminoterminales de las cadenas alfa y beta, respectivamente. Las regiones de la cadena alfa que quedan así al descubierto constituyen los "botones" que interaccionan con "ojales" del dominio terminal, y que unen los monómeros de fibrina. La unión entre monómeros se produce con un solapamiento de medio monómero (*c*). El conjunto se extiende y forma un "polímero intermediario", de dos moléculas de grosor (*d*), que se estabiliza por la formación de entrecruzamientos covalentes entre dominios terminales adyacentes. Los polímeros intermediarios se imbrican lateralmente y producen una hebra de fibrina totalmente desarrollada (*e*). La polimerización inicial parece depender específicamente de los botones alfa. El entrecruzamiento covalente, extremo con extremo, se da entre cadenas gamma. El crecimiento lateral quizá dependa de los botones beta, reforzado por entrecruzamientos de las alfa.

fácilmente. En condiciones fisiológicas, a diferencia de las experimentales, las fibras de fibrina se refuerzan y el gel se estabiliza, mediante la introducción de uniones covalentes entre monómeros adyacentes. Tales uniones se establecen gracias a la acción de un enzima, denominado factor XIII, que a su vez se activa por liberación de un péptido (catalizada por la trombina) a partir de un precursor existente en el plasma y en las plaquetas. El enzima activo une monómeros vecinos en el polímero de fibrina gracias a la formación de un enlace peptídico (el mismo tipo de enlace que une los aminoácidos en la cadena proteica) entre las cadenas laterales de lisinas (dadoras) y glutaminas (aceptoras) entrecruzándose de esta forma dos cadenas. La fibrina estabilizada por enlaces covalentes es más firme, desde el punto de vista mecánico, que la fibrina en la que no se ha producido aún el entrecruzamiento, y es también más resistente a la disolución enzimática. Sin embargo, al microscopio electrónico, no se aprecian diferencias morfológicas, por lo que el entrecruzamiento no debe introducir cambios estructurales importantes.

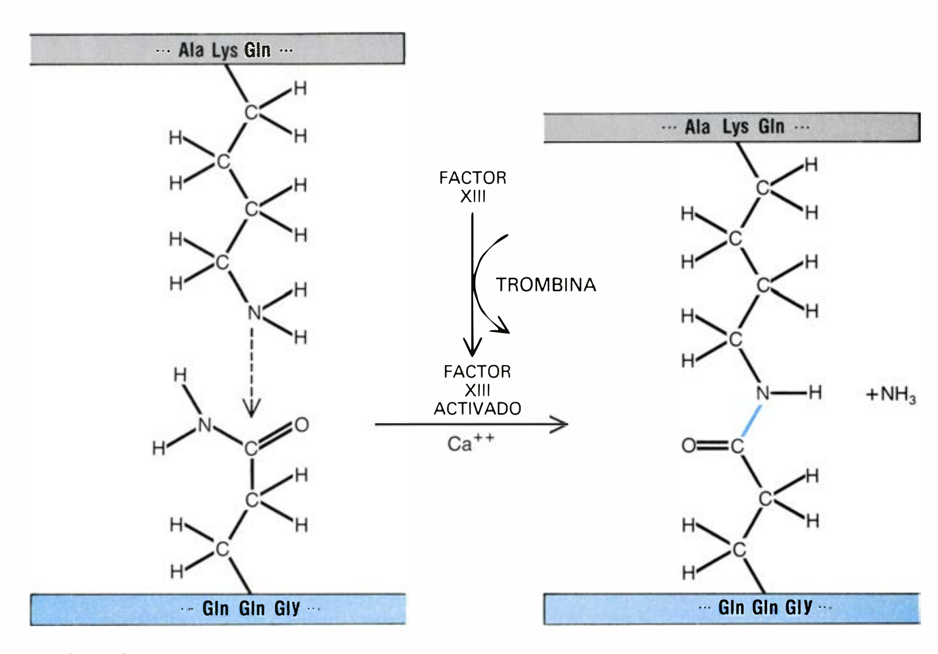
En colaboración con otro alumno mío, Renné Chen, observamos que los primeros entrecruzamientos introducidos en la fibrina durante el proceso de estabilización se daban entre cadenas gamma de dos moléculas adyacentes. Posteriormente, Patrick A. McKee, del Hospital Clínico de la Universidad de Duke, demostró que también se establecían múltiples entrecruzamientos entre cadenas alfa.

La estructura completa del entrecruzamiento gamma-gamma se estableció rápidamente. Tanto el dador (lisina) como el aceptor (glutamina) se encuentran cerca del extremo carboxiterminal de la cadena gamma. Las moléculas unidas entre sí están orientadas de forma antiparalela y se enlazan por entrecruzamientos separados por ocho aminoácidos. Puesto que todas las cadenas gamma se aparean de forma tan restrictiva, cabe suponer que todas las unidades del polímero de fibrina tienen una orientación equivalente. Las regiones carboxiterminales de la cadena gamma se encuentran en los dominios terminales del fibrinógeno; y así, la digestión, por parte de la plasmina, de la fibrina unida a las cadenas gamma origina un dímero del fragmento *D*, como se observó en el laboratorio de McKee.

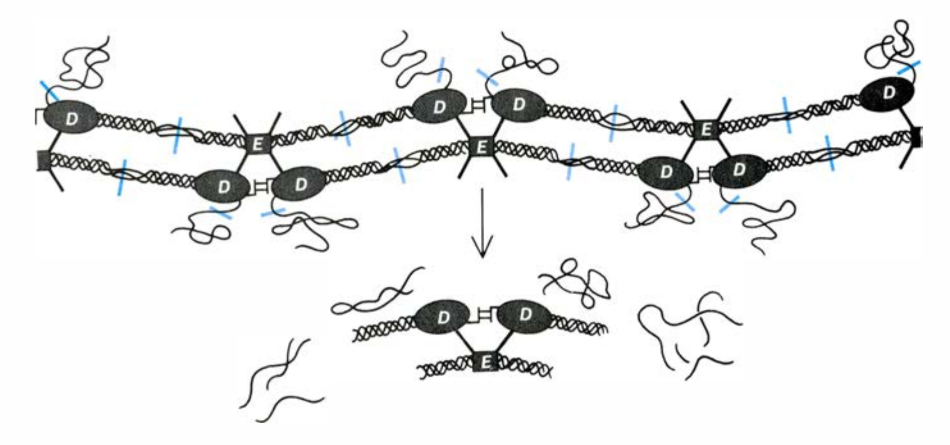
Un entrecruzamiento de cadenas gamma conecta dos monómeros para formar un dímero, mientras que las ca-



EL COAGULO DE FIBRINA se estabiliza por entrecruzamiento de dos regiones antiparalelas de las cadenas gamma. Para ello se establecen enlaces covalentes recíprocos entre dos pares lisina-glutamina.



UNION COVALENTE (C-N), formada entre lisina y glutamina con liberación de amonio (NH_3). La reacción está catalizada por un enzima dependiente del calcio, el factor XIII, activado por la trombina.



DISOLUCION DE UN COAGULO por el ataque de la plasmina. Algunos de los principales puntos de hidrólisis se muestran con líneas coloreadas. El enzima tiene fácil acceso a las regiones de conexión superenrolladas, que sufren el ataque, en primer lugar, en su región central, relativamente desestructurada. La digestión, por la plasmina, de la fibrina entrecruzada da lugar a fragmentos residuales "D_2 *E*".

denas alfa pueden entrecruzarse dando lugar a polímeros de mayor tamaño. Ello significa que cada cadena alfa está unida, al menos, a otras dos cadenas alfa, de modo que debe haber, al menos, dos glutaminas aceptoras diferentes y dos lisinas potencialmente dadoras en cada cadena alfa. Las glutaminas se localizaron con precisión durante nuestros estudios de secuenciación. En el fibrinógeno humano se encuentran separadas por 38 aminoácidos en la parte media de la cadena alfa. Las lisinas que intervienen en el entrecruzamiento de las cadenas alfa se hallan en un bloque de 200 aminoácidos, más próximas al extremo carboxilo de la cadena.

Tal como indicamos al comienzo del artículo, debe existir un mecanismo para disolver los coágulos que puedan resultar perjudiciales, bien porque se hayan liberado a la circulación general, o porque se formen en un momento o lugar inadecuados. La arquitectura de la fibrina resulta tan adecuada para su lisis como para su polimerización, si no más. La base estructural resulta evidente: la accesibilidad y vulnerabilidad de conexiones esenciales. Al limitar los contactos intermoleculares a los cuerpos nodulares de los monómeros de fibrina quedan canales abiertos hacia las conexiones internodulares, a través de los cuales pueden difundir los agentes proteolíticos (incluida la plasmina, su precursor, el plasminógeno y los activadores que convierten el precursor en enzima activo). Además, en vez de precisarse un corte a través de todo el espesor de una molécula en forma de bastón, la plasmina tiene que romper únicamente conexiones constituidas por tres hebras muy finas. En consecuencia, la disolución de la fibrina requiere sólo la hidrólisis de un pequeño número de enlaces. Puede admitirse que una sección de fibra se solubiliza por la hidrólisis de dos conexiones separadas por unos 10 a 20 monómeros de fibrina. El sistema queda reducido al estadio de "polímero intermediario", que supone la vuelta a formas intermedias producidas en la constitución del gel. La digestión proteolítica continúa hasta que los polímeros intermediarios quedan reducidos a fragmentos nucleares.

El modelo de disolución que se ha propuesto subraya tanto la naturaleza abierta de los coágulos de fibrina como la estructura, en múltiples dominios, de los monómeros de fibrinógeno y fibrina. Los productos generados por la plasmina en la degradación de la fibrina son muy similares a los producidos en la degradación experimental del fibrinógeno. Evidentemente, las asociaciones intermoleculares que enlazan entre sí las unidades de fibrina no las protegen del ataque proteolítico. Ello contrasta con muchos otros sistemas biológicos, donde la interacción de dos moléculas de gran tamaño confiere a una de ellas, o a ambas, cierta protección frente a la degradación proteolítica. La semejanza entre los productos de digestión implica también que no tiene lugar una reorganización importante durante la polimerización.

La fibrina estabilizada por uniones covalentes se hidroliza de un modo más lento que la fibrina no estabilizada, como ha demostrado Johannes Gormsen, de la Universidad de Copenhague. Ello no obstante, tras una digestión completa por plasmina, la única diferencia en los productos finales es la presencia del dímero "doble *D*" en preparaciones entrecruzadas. En condiciones menos drásticas, sin embargo, Gilbert Hudry-Clergeon, del laboratorio de Michel Suscillon, de la Universidad de Grenoble, aisló el complejo denominado "D_2 E" a partir de digeridos parciales de fibrina entrecruzada, obtenidos con plasmina. El complejo fue luego identificado y caracterizado por otros investigadores y su existencia se acepta hoy mayoritariamente. Está compuesto por una "doble *D*" y un fragmento *E*, presumiblemente de otro monómero al que los dos *D* estaban unidos antes de la digestión. El complejo "D_2 E" es exactamente lo que cabe esperar de un esquema de interacción que implique la existencia de botones en un dominio y ojales en otro.

He tratado de mostrar cómo el conocimiento de la secuencia de aminoácidos del fibrinógeno apoya algunas de las nociones generales sobre su estructura tridimensional y comportamiento. Los datos de secuencia revelan la existencia de dos dominios terminales conectados con una región central mediante haces de tres hebras, configurando una estructura trinodular semejante a la observada con el microscopio electrónico hace unos 20 años. Dicha estructura, basada en múltiples dominios, resulta muy adecuada para una serie consecutiva de fenómenos (polimerización, estabilización y fibrinolisis) que, primero, detienen la hemorragia y luego disuelven el coágulo para evitar el bloqueo de la circulación. La naturaleza molecular de esos fenómenos, puesta ya al descubierto, ha de servirnos de trampolín para llegar a dominar la manera de mantener el sistema en equilibrio en aquellos enfermos cuya sangre tiende a coagularse cuando y donde no debe.

Algebra por ordenador

Los ordenadores manipulan símbolos y números. Nuevos algoritmos de aplicación general pueden encargarse de gran variedad de tareas matemáticas rutinarias y resolver problemas hasta hoy inabordables

Richard Pavelle, Michael Rothstein y John Fitch

Entre todas las tareas que pudieran encomendarse a un ordenador electrónico, ninguna hace temer tantas dificultades como la manipulación algebraica de expresiones matemáticas complejas. Los ordenadores se han labrado ya una sólida reputación por su capacidad para ejecutar cálculos numéricos, aliviando mucho la carga que éstos imponían a los humanos. Es menos conocido, en cambio, que existan programas de ordenador igualmente bien adaptados a la manipulación de expresiones algebraicas, lo que significa que estas máquinas no sólo sean capaces de operar directamente sobre números, sino también con símbolos abstractos que los representen.

Quizás una de las razones de que la capacidad algebraica del ordenador no se haya explotado plenamente sea que la programación del ordenador es, por su naturaleza, muy parecida al álgebra. Podría pues pensarse que existe una división natural de funciones: el programador manipula símbolos algebraicos, mientras que la computadora está confinada al cálculo aritmético, a "triturar los números". Mas la falsedad de esta dicotomía fue ya reconocida en 1844 por Augusta Ada Byron, Condesa de Lovelace, e hija del poeta Lord Byron, quien fue seguramente la primera persona que programó un ordenador. Lady Lovelace era protectora de Charles Babbage, y diseñó algunos de los programas de la rudimentaria computadora mecánica que Babbage llamó "ingenio de diferencias". Augusta Ada Byron hizo ya la observación de que la computadora podía "disponer y combinar sus cantidades numéricas exactamente como si fueran letras o cualesquiera otros símbolos generales, y, de hecho, sería capaz de presentar sus resultados con notación algebraica, caso de efectuarse en él las modificaciones oportunas". Análogamente, el ordenador digital moderno es una máquina de aplicación general capaz de ejecutar cualquier algoritmo, esto es, cualquier procedimiento definido por un número finito de pasos enunciados con precisión suficiente. Los algoritmos algebraicos pueden ejecutarse en ordenador tan fácilmente cuanto puedan serlo los aritméticos.

En 1973, Richard Pavelle, uno de los autores de este artículo, se vio precisado a llevar a cabo un desarrollo algebraico en teoría de relatividad general. El empeño requirió tres meses de trabajo con lápiz y papel. Al año siguiente se presentó una tarea aún más formidable, al quererse definir aquellas propiedades matemáticas que pudieran diferenciar a la teoría de relatividad general de otras teorías gravitatorias. En lugar de intentar el desarrollo manual del nuevo cálculo, Pavelle optó por construir un programa de ordenador capaz de manipular expresiones algebraicas del tipo que aparecen comúnmente en las teorías gravitatorias. El programa se redactó seguidamente en el lenguaje de ordenador de un potente sistema de programas algebraicos llamado MACSYMA, que por aquel entonces se ponía a punto en el Instituto de Tecnología de Massachusetts. Una vez instalado en MACSYMA el programa de Pavelle, se ensayó la resolución del problema de 1973. El ordenador tardó dos minutos en confirmar el resultado de un cálculo que había requerido tres meses de trabajo. Experiencias semejantes son comunes entre los usuarios de sistemas de cálculo algebraico computarizado.

Para comprender la necesidad de sistemas automáticos de cálculo y manipulación de expresiones algebraicas se requiere tener presente que muchas nociones científicas quedan formuladas a través de enunciados matemáticos de carácter general, y que estos enunciados pierden todo su significado al sustituirlos por su evaluación numérica. Fijémonos, por ejemplo, en la sencilla expresión $3\pi^2/\pi$. Como cualquier alumno de bachillerato sabe, podemos simplificar la fracción cancelando π en el numerador y denominador, obteniendo así la expresión 3π. Puede que el valor numérico de 3π tenga interés en algunos casos, pero casi siempre será suficiente –y sin duda más útil– dejar la expresión en forma literal, simbólica, no numérica. Con un ordenador programado exclusivamente para el cálculo aritmético es forzoso calcular el valor de la expresión $3\pi^2/\pi$; si el cálculo se hace con precisión de hasta 10 cifras significativas el valor obtenido será 9,424777958. Este número, sobre ser una cadena de dígitos que nada nos dice, es distinto del calculado al determinar el valor numérico (con diez cifras decimales) de 3π: este último vale 9,424777962. La discrepancia de las dos últimas cifras decimales procede de errores de redondeo y truncamiento, introducidos por el ordenador. En un ordenador programado así no es probable que los números $3\pi^2/\pi$ y 3π se consideren equivalentes.

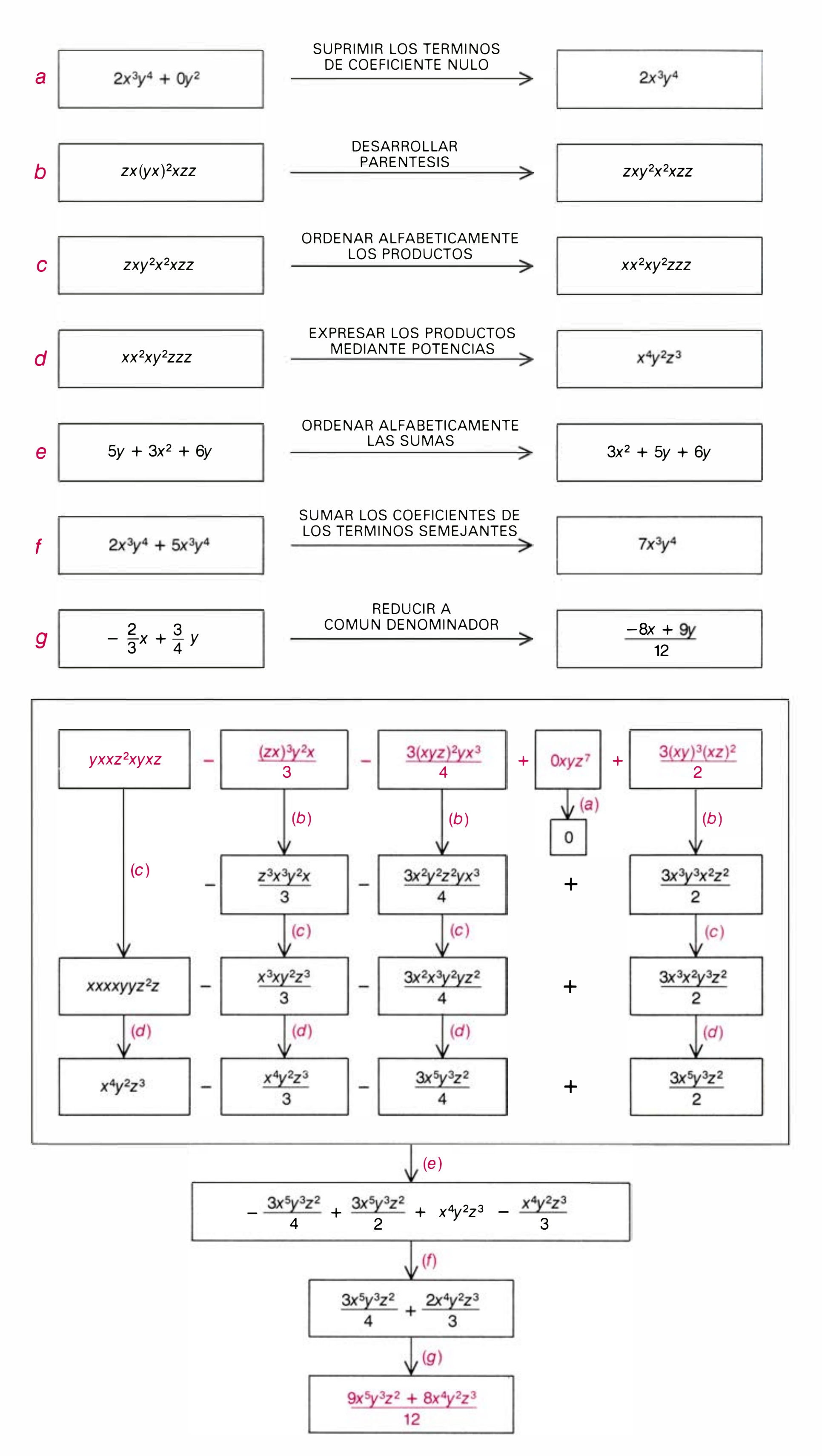
El ejemplo expuesto pone de relieve tres ventajas que ofrecen los programas algebraicos sobre los numéricos. En primer lugar, con frecuencia es más económico (en tiempo de funcionamiento del ordenador) simplificar las expresiones algebraicas antes de calcular sus valores numéricos. Aunque en el ejemplo anterior el ahorro de tiempo sea trivial, hay muchos problemas complicados donde la economía producida por simplificación algebraica no es despreciable. Segundo, a diferencia de las aproximaciones numéricas generadas por el ordenador, las soluciones algebraicas son exactas. En el cálculo aproximado es inevitable cometer errores;

cuando han de efectuarse muchas operaciones sucesivas con números aproximados, los errores pueden ir acumulándose y provocar resultados finales absurdos. Tan sólo tras cuidadoso análisis del error podrá darse por válida la solución final. Y el análisis del error es con frecuencia uno de los problemas más complejos que deben afrontarse en muchas parcelas de la ciencia.

La tercera –y tal vez la más importante ventaja– es que suele ocurrir que los fines de una investigación queden mejor servidos presentando los resultados en forma algebraica. Como bien dice Richard W. Hamming, de los Laboratorios Bell, "la finalidad del cálculo no es darnos números, sino ideas". En ocasiones podemos hacernos con tales ideas calculando valores en una expresión matemática; pero la mayoría de las veces, las relaciones entre magnitudes se ven más claramente expresándolas en forma algebraica.

Así, por ejemplo, al estudiar un proceso químico se ha logrado expresar matemáticamente la relación entre la estabilidad del proceso y las cantidades relativas o concentraciones de las sustancias allí presentes. A partir de tales relaciones podemos predecir con exactitud suficiente si una pequeña modificación de alguna de las magnitudes tendrá por efecto una reacción violenta o controlada. Análogamente, en la teoría de la evolución estelar es posible examinar algebraicamente de qué manera cierto número de variables determinan si la estrella se convertirá en estrella de neutrones o en un "agujero negro", o cómo interactúan las variables para predecir la existencia de un nuevo objeto.

Los sistemas automáticos que vamos a denominar "sistemas algebraicos computarizados" presentan una cuarta ventaja, que salta a la vista tan pronto se examinan ciertos aspectos prácticos de la acumulación de conocimientos científicos. No son raros los casos donde las teorías científicas se enuncian mediante expresiones matemáticas concisas y sumamente generales que sugieren la forma que adoptará la teoría bajo ciertas hipótesis. Por ejemplo, cierta parcela de la teoría de relatividad general, la que estudia las ecuaciones einstenianas del campo, puede expresarse en una sola línea, admitidos ciertos convenios notacionales de uso común. Explorar las implicaciones de carácter físico de la teoría pasa por la exploración matemática de los hechos matemáticos implícitos en las ecuacio-



ALGORITMO DE ORDENADOR para simplificar expresiones algebraicas por reducción de términos semejantes. Este algoritmo recuerda los métodos ordinarios de simplificación manual. Es preciso detallar cuidadosamente al ordenador cada etapa del proceso. En el diagrama superior se ilustra el efecto de cada paso sobre una expresión algebraica sencilla. En la parte inferior se muestra, mediante un organigrama, la forma de aplicar el algoritmo a una expresión más complicada. Las letras rojas situadas junto a las flechas del organigrama corresponden a las operaciones elementales del diagrama superior. La fórmula que el algoritmo genera a partir de la expresión figura en rojo en el recuadro inferior. Una de las tareas que ha de ejecutar cualquier sistema de programas de ordenador que pretenda resolver problemas matemáticos dados por símbolos es la simplificación de expresiones. Aunque el algoritmo de simplificación genera una expresión útil, no existe ninguno capaz de reducir toda expresión a su forma más sencilla.

```
(C1) F(J):=FACTOR(NUSUM(K↑J,K,O,N));

(D1) F(J) := FACTOR(NUSUM(K^J, K, O, N))

(C2) FOR J:1 THRU 2 DO DISPLAY (SUM(K↑J ,K,O,N) = F(J));

(D2)  N
     ====
     \          N (N + 1)
      >     K = ---------
     /              2
     ====
     K = O

      N
     ====
     \           N (N + 1) (2 N + 1)
      >     K^2 = -------------------
     /                    6
     ====
     K = O

(C1) (X+7)↑10;

(D1) (X + 7)^10

(C2) EXPAND(D1);

(D2)           X^10 +       70 X^9 +      2205 X^8 +     41160 X^7 + 504210 X^6 + 4235364 X^5

      + 24706290 X^4  + 98825160 X^3 + 259416045 X^2 + 403536070 X  + 282475249

(C1) X↑3+B*X↑2+A↑2*X↑2-9*A*X↑2+A↑2*B*X-2*A*B*X-9*A↑3*X+14*A↑2*X-2*A↑3*B+14*A↑4=0;

(D1) X^3 + B X^2 + A^2 X^2 - 9 A X^2 + A^2 B X - 2 A B X

      - 9 A^3 X + 14 A^2 X - 2 A^3 B + 14 A^4 = 0

(C2) SOLVE(D1,X);

(D2) [X = 7 A - B, X = - A^2, X = 2 A]

(D1)            1125899906842624 Y^40 +       9007199254740992 Y^39 +       43523068273885184 Y^38

          + 96686654500110336 Y^37 +       71892942171668480 Y^36 -       203545990580404224 Y^35

           - 3231739551940608 Y^34 +     2967153761027358720 Y^33 +     3933037175129505792 Y^32

       - 10392801849559220224 Y^31 -     26535490294760079360 Y^30 +     60970867840763559936 Y^29

      + 124971270109665427456 Y^28 -    294061554052926275584 Y^27 -    377892103185282105344 Y^26

     + 1043007763435371888640 Y^25 +  1247934126027765186560 Y^24 -   4589204662544561602560 Y^23

       - 5344792849977515704 32 Y^22 + 13715332431765689073664 Y^21 - 13302644475428239835136 Y^20

     - 11542363148164431609856 Y^19 + 30184905793343524306944 Y^18 - 11553929011986625003520 Y^17

     - 20761961484788844621824 Y^16 + 23716893785532021784576 Y^15 -  14706887769816009359 36 Y^14

     - 11829726550983847370752 Y^13 +  7047410412446462752768 Y^12 +   5906483242725501337 60 Y^11

      - 2234695017099712874496 Y^10 +   836560570027035444224 Y^9  +     38308661311225029504 Y^8

       - 113171442907550345472 Y^7  +     29418059499453088640 Y^6  +        66680780706292864 Y^5

         - 1434405693611092512 Y^4  +        28386128184561 0304 Y^3  -        20855579198349088 Y^2

             + 370175331659648 Y   - 1886186974063

(C2) FACTOR(D1);

Time= 130855 msec.

(D2)   (4096 Y^10 +    8192 Y^9 -    3008 Y^8 -   30848 Y^7 +  21056 Y^6 + 146496 Y^5

     - 221360 Y^4  +    1232 Y^3 + 144464 Y^2 -   78488 Y  + 11993)

       *(4096 Y^10 +    8192 Y^9 +    1600 Y^8 -   20608 Y^7 +  20032 Y^6 + 87360 Y^5

     - 105904 Y^4  +   18544 Y^3 +  11888 Y^2 -    3416 Y  + 41)

       *(8192 Y^10 +   12288 Y^9 +   66560 Y^8 -   22528 Y^7 - 138240 Y^6 + 572928 Y^5

      - 90496 Y^4  -  356032 Y^3 + 113032 Y^2 +   23420 Y  - 8179)

       *(8192 Y^10 +   20480 Y^9 +   58368 Y^8 -  161792 Y^7 + 198656 Y^6 + 199680 Y^5

     - 414848 Y^4  -    4160 Y^3 + 171816 Y^2 -   48556 Y  + 469)
```

SISTEMA DE ALGEBRA por ordenador llamado MACSYMA, puesto a punto por el Instituto de Tecnología de Massachusetts. Es un conjunto de programas de aplicación general, capaz de resolver problemas algebraicos que superan con mucho la capacidad de la mayoría de los matemáticos. El sistema se comunica con el usuario, y mantiene con él una especie de diálogo. Una vez puesto en comunicación con el sistema, el usuario dactilografía la expresión algebraica a manipular, empleando un lenguaje de alto nivel semejante al FORTRAN o el ALGOL, que escribe en la línea *C* (línea de mando). El sistema responde presentando en la línea *D* una versión escrita con notación semejante a la habitual en matemáticas. El usuario da una segunda instrucción a través del teclado, y el ordenador ejecuta la operación que se le ha encargado. En los dos primeros ejemplos, el ordenador halla una expresión de la suma de los *n* primeros términos de una serie, cuyos sumandos son todos de la forma inmediata al signo de sumación Σ. (Σk^2 denota la suma de $1^2 + 2^2 + 3^2 + \ldots + n^2$.) En los dos ejemplos siguientes, el ordenador desarrolla la potencia de un binomio y resuelve exactamente una ecuación cúbica. Las flechas que apuntan hacia arriba denotan exponenciación y, los asteriscos, multiplicación. El último de los ejemplos mostrados nos da una idea de la potencia de cálculo del sistema. El área máxima de un hexágono inscrito en un círculo de diámetro 1 viene dada por una de las raíces del enorme polinomio de grado 40 que vemos en el centro. Los métodos manuales ordinarios quedan impotentes frente a polinomios de tan gran tamaño; es más fácil descomponer el polinomio en factores y hallar las raíces de éstos. El sistema de álgebra computarizada proporciona la factorización más completa que pueda lograrse con polinomios de coeficientes enteros.

nes de campo; ahora bien, salvo escasas excepciones, tal exploración matemática requiere efectuar operaciones algebraicas tan complejas que su tratamiento "a mano" resulta impracticable. Aun admitiendo hipótesis que simplifican enormemente el tratamiento del problema, la tarea algebraica es tan penosa que pocos investigadores están dispuestos a arriesgar su salud, su felicidad y su bienestar profesional tratando de alcanzar una solución.

Los resultados de tales trabajos, aun cuando lleguen a ser admitidos en el cuerpo principal del saber científico, son como puestos avanzados, aislados, raramente comprobados por otros investigadores, y, por lo común, aceptados sin más fundamento que la reputación de su autor y no como resultado de minucioso examen crítico. Tan lamentable estado de cosas puede ocurrir en cualquier rama de la ciencia cuando nos aproximamos a lo que podríamos llamar "la frontera de lo inabordable". Hoy, la posibilidad de efectuar cálculos algebraicos computarizados ha dilatado significativamente la región incluida entre fronteras.

En mecánica celeste, la estirpe de cálculos algebraicos extremadamente largos tiene ya distinguida tradición. En 1847, el astrónomo francés Charles Delaunay empezó a calcular la posición de la Luna en función del tiempo. En cierto sentido, tal cálculo era aplicación directa de la teoría newtoniana de la gravitación universal. En la teoría newtoniana, la órbita de una masa puntual en torno a una masa esférica de densidad uniforme es una elipse, pero las características del sistema Tierra-Luna determinan que la órbita lunar sea considerablemente más compleja que la elipse. El plano de la órbita lunar está inclinado, formando un pequeño ángulo con respecto al plano de la órbita terrestre en torno al Sol, ángulo que varía a causa de la perturbadora influencia del campo gravitatorio solar. También el Sol es causa de que el perigeo lunar (punto de la órbita lunar situado a mínima distancia de la Tierra) sufra, con respecto a las estrellas, un lento movimiento de precesión.

En vista de tales complicaciones no se solía determinar la órbita lunar mediante una función del tiempo, como se propuso Delaunay, sino mediante extrapolación numérica basada en el conocimiento de posiciones previas de nuestro satélite. En períodos cortos, los errores acumulados en la extrapolación son pequeños, y pueden corregirse de-

terminando nuevamente la posición lunar, por observación. Lograr cálculos lo bastante rápidos como para marchar al paso del cambio de posición lunar es francamente difícil, incluso con los ordenadores más rápidos; además, las observaciones requeridas para conservar el ajuste y la exactitud de las previsiones son costosas y requieren tiempo. El cálculo de Delaunay consiguió remontar estas deficiencias, pero a costa del más extremo sacrificio, pues tuvo que dedicar 20 años a realizarlo y comprobarlo.

Los resultados del trabajo de Delaunay se publicaron en 1867, en dos volúmenes, aceptándose sin reparos durante más de 100 años. Durante la segunda guerra mundial renació el interés por expresar las coordenadas de estrellas fijas en función del tiempo, con el fin de facilitar la localización de los submarinos alemanes, pero no se realizaron nuevos trabajos acerca de la órbita lunar o de otros cuerpos del sistema solar. Con el advenimiento de los satélites artificiales, el método de Delaunay adquirió, definitivamente, interés comercial.

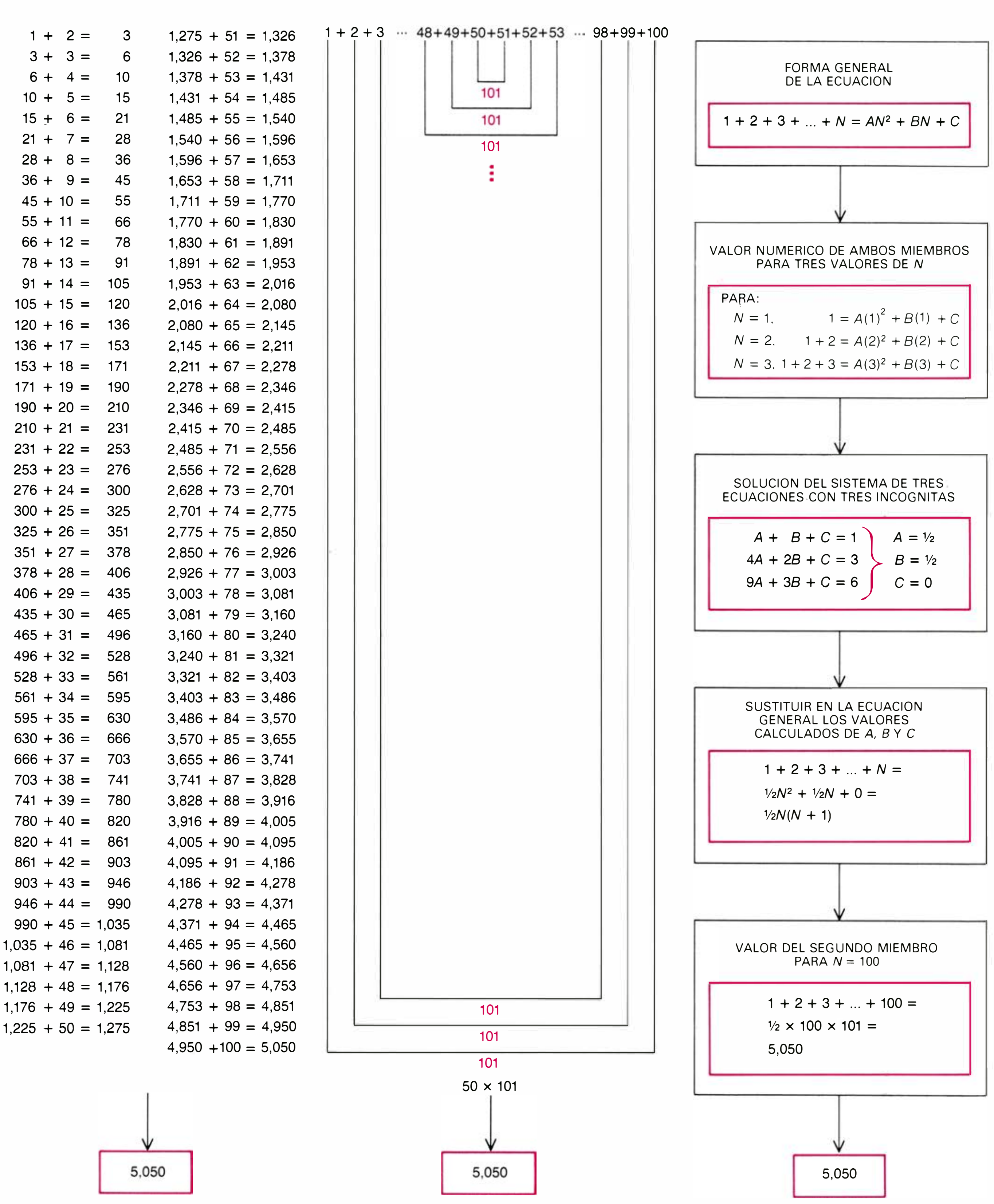
Para preparar expresiones algebraicas de las órbitas de los satélites, André Deprit, Jacques Henrard y Arnold Rom, a la sazón en los Boeing Scientific Research Laboratories de Seattle, se pusieron a investigar sobre la construcción de algoritmos capaces de ejecutar los cálculos de Delaunay con auxilio de un ordenador. En 1970 consiguieron efectuar el cálculo automáticamente, en 20 horas. Vale la pena destacar que solamente se encontraron tres errores en el trabajo de Delaunay, todos en términos de pequeño valor, y donde, además, dos de ellos eran meras consecuencias del otro error. Desde entonces se han desarrollado docenas de sistemas de álgebra computarizada, que han generalizado el método de Delaunay hasta el extremo de tomar en cuenta los efectos del rozamiento con las altas capas de la atmósfera y la no esfericidad de la Tierra, que perturban las órbitas de satélites de baja altura. Tales algoritmos se emplean hoy para seguimiento de estos ingenios.

Habida cuenta de la duración de la vida humana, lograr que el tiempo requerido para una tarea pase de 20 años a 20 minutos comporta una ganancia cualitativa. Cabe así en lo posible que problemas tan complejos como el de Delaunay atraigan la atención de investigadores mucho menos pacientes. Además, el investigador puede ahora dedicar una proporción mucho mayor de su tiempo a sopesar el significado e importancia de sus descubrimientos algebraicos. Puede experimentar cálculos sumamente prolijos sin apostar su carrera a la corrección y utilidad de su solución. Quizá la indicación más expresiva de que el álgebra computarizada ha alejado muchísimo las fronteras de la intractabilidad en mecánica celeste sea que el enorme cálculo de Delaunay se emplea hoy para ensayar los nuevos sistemas algebraicos computarizados.

¿Cómo se programa un ordenador para que efectúe manipulaciones algebraicas? Muchos de los algoritmos empleados en el álgebra computarizada empezaron siendo procedimientos básicos ideados por anteriores generaciones de matemáticos; hoy se enseñan a los alumnos de bachillerato y en los primeros cursos de universidad. Por ejemplo, los métodos de resolución de ecuaciones que contienen una sola cantidad desconocida elevada a las potencias primera o segunda datan de los tiempos de Hammurabi, hacia 1750 a. de C. Pero, a diferencia de los estudiantes, que pueden ir asimilando de forma intuitiva y más o menos vaga cómo deben proceder para lograr la solución, el ordenador ha de seguir una rutina, un protocolo riguroso, que en cada momento especifique lo que debe hacer.

Fijémonos cómo procedería un principiante para resolver la ecuación $3x-1=2\ (x+5)$. Seguramente empezaría "eliminando paréntesis", multiplicando por 2 los dos sumandos x y 5, con lo que obtendría $3x-1=2x+10$. Seguidamente transpondría el término $2x$ al primer miembro de la ecuación, y el sumando -1 al segundo, cambiándoles el signo en el proceso, y reduciendo (sumando o restando) los términos semejantes. Con un último paso obtendría la solución $x = 11$.

En un programa de álgebra computarizada los pasos a seguir para resolver un problema tienen que especificarse hasta un grado de detalle bastante tedioso, y los atajos y simplificaciones que el estudiante pudiera apreciar al trabajar en un problema concreto han de detallarse cuidadosamente. El primer paso del programa para resolver la ecuación seguramente no sería distinto del que realizó nuestro estudiante: el programa suprimiría los paréntesis, obteniendo la expresión $3x-1=2x+10$. Sin embargo, para llegar a la última línea, el programa tendría que transferir uno por uno los términos de la ecuación a un miembro conveniente. Una forma



APLICANDO ALGEBRA COMPUTARIZADA a un cálculo numérico puede economizarse tiempo de procesado, por simplificación de expresiones algebraicas antes de su evaluación numérica. Para hallar directamente la suma de los 100 primeros números enteros hay que hacer 99 sumas. Un algoritmo más sencillo, empleado por el programa de álgebra computarizada, resuelve este problema valiéndose de que un resultado matemático general establece que la suma de los n primeros números enteros ha de ser función cuadrática (polinómica de segundo grado) de n. Análogamente, el algoritmo calcularía la suma de los cuadrados de los n primeros enteros construyendo una función polinómica de tercer grado en n. Aunque este algoritmo es eficaz, no es necesariamente el método más rápido para resolver el problema, ni tampoco el que proporciona más clara comprensión. Según la leyenda, cuando el matemático alemán Carl Friedrich Gauss contaba siete años observó que los enteros de 1 a 100 podían agruparse por pares con la misma suma, 101. La suma de los 100 números es, por tanto, igual a 101 multiplicado por 50, que es el número de pares. Por lo común, los sistemas de álgebra computarizada no son capaces de observaciones semejantes; por otra parte, la perspicaz idea de Gauss no es aplicable a sumas de cuadrados, cubos, o mayores potencias de los enteros.

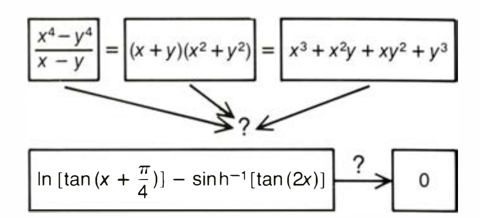
de lograrlo sería programar el ordenador para que sume a cada miembro el opuesto (aditivo) de todas las constantes que figuren en el primer miembro (o sea, el lado izquierdo de signo igual) y de todos los términos literales (esto es, que contengan incógnitas) del segundo miembro. Como la única constante que figura en el primer miembro es -1, el ordenador sumaría a toda la ecuación su opuesto, $+1$, obteniendo $3x-1+1=2x+10+1$. Seguidamente reordenaría los términos, de ser necesario, y sumaría y restaría donde fuera posible, obteniendo $3x = 2x + 11$.

A continuación, el programa examinaría las expresiones del segundo miembro en búsqueda de términos que contuvieran incógnitas, y determinaría el opuesto aditivo del primero de tales términos con que se tropezase. En nuestro ejemplo, el ordenador sumaría a los dos miembros el opuesto de $2x$, que es $-2x$, obteniendo ahora $3x-2x=2x+11-2x$. Nuevamente sería preciso reordenar los términos para reducir los que fueran semejantes: $3x-2x=2x-2x+11$. Finalmente, el programa reduciría los términos que contuvieran incógnitas de nombres iguales, terminando cuando en el primer miembro de la ecuación figurase una sola incógnita y en el segundo una única constante numérica: $x = 11$. Semejante algoritmo no es ni el más sencillo ni tampoco el más eficiente que pueda idearse para la programación de ordenadores con el objeto de resolver ecuaciones de una incógnita; por otra parte, tampoco está lo bastante elaborado. Hay muchas posibilidades; por ejemplo, la aparición de expresiones como $2x = 22$, frente a las cuales este algoritmo no sabría qué hacer.

Al diseñar un algoritmo para el tratamiento computarizado de expresiones algebraicas, no es forzoso ajustarse a los procedimientos más eficaces del cálculo manual. Delaunay demostró varios teoremas que aplicó seguidamente para simplificar cálculos intermedios. Aunque podían incorporarse tales teoremas a un programa, el esfuerzo necesario para expresarlos algorítmicamente animó a Deprit y sus colegas a buscar un método más compatible y adaptado a la ejecución mecánica del cálculo. El método que finalmente inventaron requiere transformaciones que hubieran desbordado incluso la capacidad de un Delaunay, pero su algoritmo es fácil de programar y de rápida ejecución en el ordenador. La creación y puesta a punto de nuevos algoritmos constituye una de las parcelas de investigación más activas en álgebra computarizada; y es sobre todo a raíz de estos trabajos por lo que los sistemas de álgebra automática han experimentado importantes perfeccionamientos durante estos últimos años.

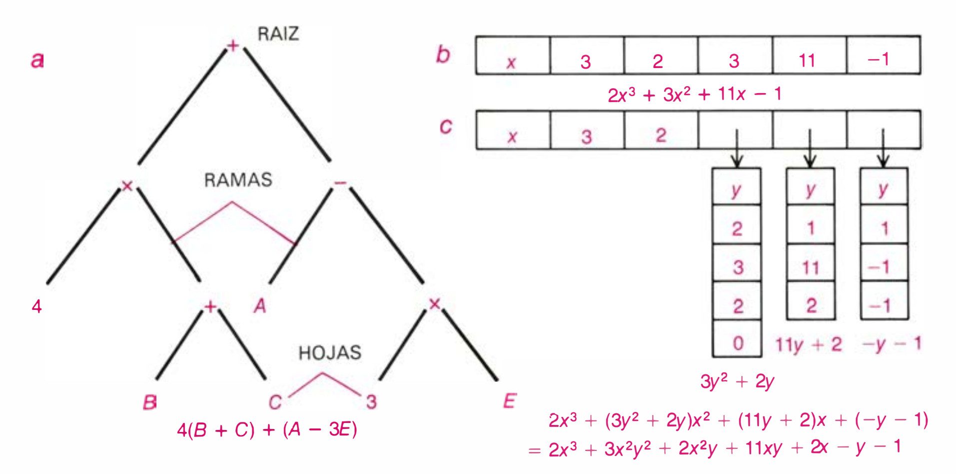
Al representar expresiones algebraicas mediante programas de ordenador, en la mayoría de los sistemas se busca almacenar la información mínima necesaria capaz de especificar unívocamente cada expresión. Los actuales esquemas de representación se sirven de una de dos técnicas básicas, o de combinaciones de ambas. En una de éstas, la expresión se representa mediante una estructura "arborescente" invertida, en la cual las "hojas" serían los operandos. Supongamos, por ejemplo, que deba representarse mediante ordenador la fórmula $2(x + 4)$. Las hojas del árbol serían los términos 2, x y 4; si bien aparecerían a distintos niveles. Los términos x y 4 estarían conectados mediante ramas ascendentes a un signo "+". A su vez, del 2 y del signo + partirían ramas ascendentes que se reunirían en lo alto –la "raíz"– del árbol, ocupada por un signo de multiplicación. La representación por este esquema hace más eficiente la búsqueda de expresiones subordinadas de formas preestablecidas.

En el segundo esquema de representación se van asignando, en orden definido, "casillas" (*slots*) donde represen-



AL SIMPLIFICAR una expresión, el sistema de álgebra computarizada tropieza con el problema de elegir, entre las diversas formas equivalentes de una misma expresión, como las que vemos arriba, cuál habrá de ser considerada como más sencilla. Muchos de los sistemas en fase de diseño dejan parte de tal decisión a la persona que interactúa con ellos. La compleja expresión de la segunda línea es igual a cero para todos aquellos valores de *x* que den valor real y positivo a las tangentes. Sin embargo, la expresión no es idénticamente nula para todos los valores de *x*. Si el sistema de cálculo computarizado sustituyera directamente la segunda expresión por cero, quedaría muy oscurecida la relación entre la expresión inicial y el resultado de la simplificación. Las expresiones *ln*, *tan* y *sinh*$^{-1}$ denotan, respectivamente, el logaritmo natural, la tangente trigonométrica y el argumento seno hiperbólico, función inversa del seno hiperbólico.

tar la información contenida en una expresión. Por ejemplo, para representar un polinomio de una variable se asigna una casilla al nombre de la variable, la siguiente al grado del polinomio (el máximo exponente a que figure elevada la variable) y las posteriores, a los coeficientes de las potencias (por orden de grados decrecientes) de la variable. El sistema de casillas de información puede quedar integrado en una representación arborescente cuando sea preciso almacenar expresiones más complicadas.



LA REPRESENTACION DE FUNCIONES ALGEBRAICAS mediante ordenador puede requerir varias estrategias. El diagrama arborescente invertido (*a*) proporciona un método sencillo y adecuado para representar la secuencia de operaciones a realizar con las variables y constantes de una expresión. Una función polinómica de una variable puede representarse mediante una sucesión (*b*) formada por el nombre de la variable, el grado del polinomio (que es la máxima potencia a que aparece elevada la variable) y los valores numéricos de los coeficientes de las potencias decrecientes de la variable. Para representar polinomios de dos variables (*c*) puede emplearse una generalización de la representación de polinomios de una variable. Aquí, el grupo de casillas horizontales representa un polinomio de tercer grado de la variable *x*, mientras que las columnas de casillas representan funciones polinómicas de la variable *y*. Se sobreentiende que los tres polinomios en *y* son los coeficientes de los términos del polinomio en *x* correspondiente a los tres huecos del casillero horizontal. Por tanto, en el ejemplo, el coeficiente de x^2 es $3y^2 + 2y$, el coeficiente del término *x* es $11y + 2$ y el del término constante respecto de *x* en el polinomio es $-y - 1$.

A pesar de que en el ordenador cada representación debe determinar unívocamente lo representado, una misma expresión algebraica puede quedar representada de diversas formas mutuamente equivalentes. Tal vez quisiéramos construir un algoritmo donde cada expresión permaneciera siempre representada en la forma más sencilla posible; en tal algoritmo, por ejemplo, $x + x + x$ seguramente se representaría en la forma $3x$. No existe, sin embargo, acuerdo sobre cuál sea la forma más sencilla de una expresión. Las preferencias estilísticas se traducen en diferencias de concepción que recuerdan no poco las diferencias de ideología política: en los programas diseñados por "conservadores" las expresiones no se someten a transformación, a menos que el usuario del programa así lo especifique; en cambio, en los programas de corte "progresista" las expresiones se transforman siempre para acomodarlas a la configuración que más grata le resultaría al programador. Mucha gente desea conservar cierto control sobre la simplificación de sus expresiones, porque la utilidad de cada forma concreta puede depender del contexto en que se la encuentre. Sea como fuere, aun admitiendo que los humanos lograsen ponerse de acuerdo sobre la forma más sencilla de cada expresión, sería preciso tener en cuenta que ya está demostrada la imposibilidad de construir un algoritmo general de simplificación.

Los sistemas de álgebra computarizada son capaces de simplificar expresiones insoportablemente largas y complejas. En tales expresiones tienen cabida no sólo las funciones elementales (polinómicas, trigonométricas y logarítmicas, así como las combinaciones de ellas), sino también funciones más complejas que puedan presentarse en el trabajo científico. El usuario de un sistema de álgebra computarizada puede también definir sus propias funciones, especificar sus propiedades, y suministrar al sistema las reglas de simplificación pertinentes; más aún, en ciertos casos puede utilizarse el ordenador para descubrir tales reglas. De cualquier forma, casi nunca se sabe por adelantado nada sobre el aspecto que tendrá el resultado de la simplificación. El método más práctico parece consistir en solicitar de la máquina todos los tipos de expresiones cuya simplificación tiene ya almacenada en su memoria, y suministrar al programa cualesquiera nuevas identidades algebraicas que el programador considere pertinentes, dejando luego que el ordenador proceda por su cuenta. El método funciona sorprendentemente bien; con frecuencia, la simplificación lograda es equiparable a la de un matemático experto.

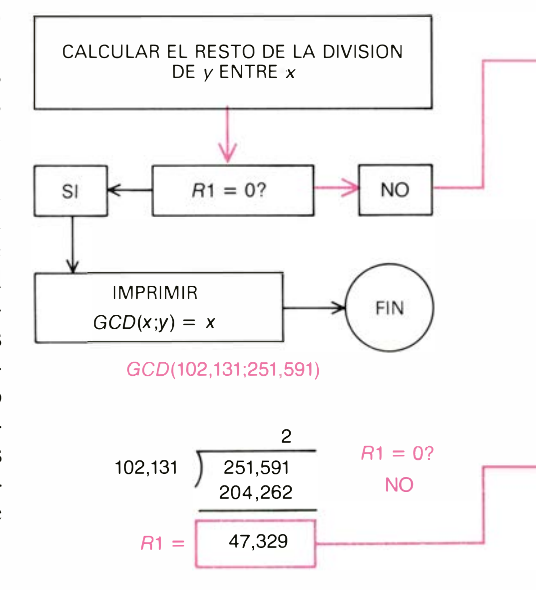
En los cálculos algebraicos es necesario que los resultados sean exactos; no son admisibles los valores aproximados. Por ello, un sistema algebraico ha de poder manejar números enormes con precisión ilimitada. En casi todos los ordenadores, los números se representan corrientemente con un número prefijado de cifras significativas, aunque, a veces, la precisión de las operaciones aritméticas puede hacerse con doble (o aún mayor) número de dígitos. Sin embargo, para las necesidades del álgebra computarizada es imposible imponer *a priori* límites de precisión.

Para comprenderlo, fijémonos en un algoritmo que puede ser útil al simplificar expresiones racionales (cocientes indicados de polinomios). Tal algoritmo resulta de generalizar y modificar el método llamado "algoritmo de Euclides", que permite calcular el máximo común divisor (m.c.d.) de dos números; tal método data del siglo III a. de C. El m.c.d. de dos números, por ejemplo 6 y 8, se halla mediante una serie fija de divisiones repetidas. Una vez hallado el m.c.d. del numerador y denominador de una fracción, ésta puede transformarse en otra equivalente e irreducible (cuyos términos son mínimos), sin más que dividir su numerador y denominador por el m.c.d. de ambos. En la fracción 6/8, el m.c.d. es 2, y la fracción reducida es 3/4.

Sin embargo, al aplicar el algoritmo de Euclides a polinomios, es fácil que los cocientes y restos de las sucesivas divisiones generen pronto términos cuyos coeficientes sean fracciones con términos enormemente grandes. Se han ideado ciertas técnicas que evitan tener que calcular con números tan grandes, reduciendo algo el tamaño de los números que van surgiendo en el proceso. Por ejemplo, si durante la ejecución del algoritmo de Euclides todos los coeficientes admiten un factor común, podemos simplificar dividiendo por él. Sin embargo, bastaría con redondear uno de los coeficientes antes de terminar el cálculo para que la respuesta final del programa careciese de sentido.

Uno de los más importantes algoritmos puestos a punto durante los quince últimos años para los sistemas de álgebra computarizada resuelve problemas de integración indefinida. El proceso de integración –o "primitivación" como algunos gustan llamarlo– es de la máxima importancia en cálculo infinitesimal; lo descubrieron, independientemente, Newton y Leibniz. El proceso de integración permite combinar un número ilimitado de cantidades arbitrariamente pequeñas, y generar una cantidad –un número– finita, bien especificada. En la llamada integración indefinida, el método es aún más general: el resultado de la integración no viene dado por un número concreto, sino como una función de una o más variables. Es frecuente tropezar con problemas de esta índole en ciencias físicas y biológicas, pero la dificultad para resolver sistemáticamente este problema es tan grande que ha desesperado a los matemáticos durante varios siglos. Ha llegado a pensarse que sería imposible lograr un algoritmo capaz de resolver el problema en su forma general; el método tradicional de realizar integraciones se basa en buenas dosis de experiencia, "ojo clínico" y el manejo de tablas de integrales ya resueltas.

El algoritmo de integración indefinida que por fin ha logrado construirse se asemeja a otro algoritmo posterior, preparado para determinar expresiones concisas de sumas de un número finito de expresiones que sean todas de formas semejantes [*véase la ilustración de la página 86*]. Si una integral puede expresarse en forma explícita, mediante combinación finita de funciones elementales (no, por tanto, si el desarrollo



ORGANIGRAMA que detalla el desarrollo del algoritmo de Euclides, un procedimiento riguroso para determinar el máximo común divisor (m.c.d., simbolizado aquí con sus iniciales inglesas GCD)

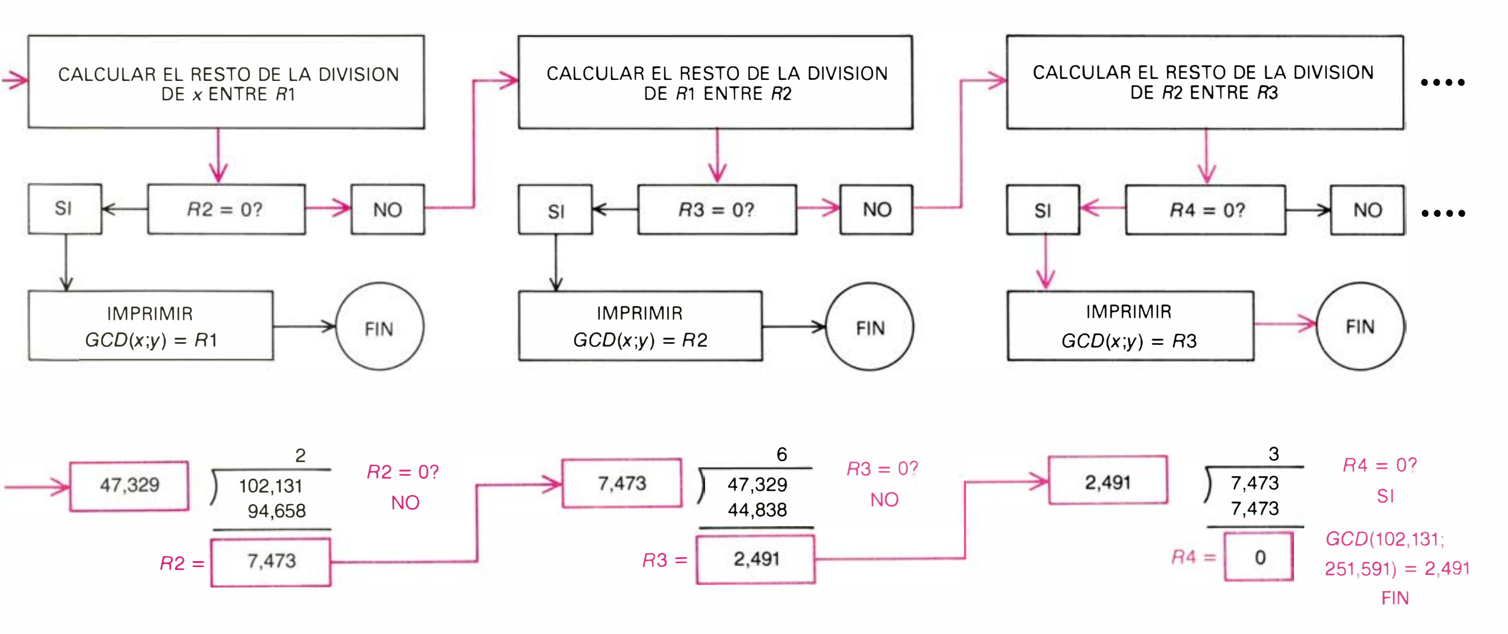
requiere una serie infinita) es posible predecir la forma algebraica general de la solución y, después, ir trabajando a la inversa, por derivación (proceso recíproco de la integración) hasta determinar la expresión exacta. Tal algoritmo ha entrado a formar parte de varios sistemas de álgebra computarizada. Un estudio numérico computarizado de ocho tablas de integrales indefinidas de amplia difusión ha permitido descubrir que alrededor del 10 por ciento de las fórmulas contienen errores; en una de las tablas, la tasa de fallos alcanzó el 25 por ciento.

Es probable que en los Estados Unidos no se inviertan más de 3 millones de dólares (unos 300 millones de pesetas) anuales en el desarrollo de sistemas de álgebra computarizada; a pesar de tan baja inversión, los sistemas existentes están ya dando prueba de gran elaboración y potencia de cálculo. No menos rica es la diversidad de sus diseños. Existen hoy unos 60 de tales sistemas, que podemos agrupar en tres categorías, atendiendo a su desarrollo histórico. Los sistemas del primer grupo son descendientes de los primeros intentos de redacción de programas de álgebra computarizada, que se diseñaron para resolver problemas específicos en campos como la física matemática o la química teórica. En estos programas "a la medida" es factible sintonizar exactamente el programa con el tipo de problemas a resolver, por cuya razón suelen operar a gran velocidad. Actualmente existen programas de este tipo para problemas de electrodinámica cuántica (ASHMEDAI), teoría lunar y relatividad general (CAMAL), física de altas energías (SCHOONSCHIP), manipulación de tensores subindicados (SHEEP), mecánica celeste (TRIGMAN) y para la resolución de ecuaciones polinómicas o entre funciones racionales (ALTRAN). Hay incluso un programa especializado, llamado ALDESSAC/2, cuya función es servir de ayuda a la preparación y puesta a punto de programas de álgebra computarizada.

Los sistemas del segundo bloque son los llamados "de aplicación general". Estos sistemas buscan proporcionar al investigador todo el instrumental matemático posible. Son capaces de ejecutar las cuatro operaciones aritméticas fundamentales, llevar a cabo los procesos de integración definida e indefinida, resolver ecuaciones algebraicas y diferenciales ordinarias (en las ecuaciones diferenciales se relaciona la tasa de variación de una magnitud con los valores y tasas de variación de otras magnitudes). Los programas algebraicos de aplicación general pueden igualmente resolver sistemas de ecuaciones algebraicas, lineales y no lineales, derivar funciones, simplificar fórmulas, descomponer en factores y calcular con desarrollos de Taylor, finitos o infinitos. Pueden, además, ejecutar todas las funciones de los programas especializados, aunque no con tanta velocidad ni potencia. Los cuatro sistemas más conocidos de aplicación general son MACSYMA, ya mencionado, REDUCE, sistema desarrollado entre las Universidades de Stanford y Utah en colaboración con la Rand Corporation; SCRATCHPAD, creado por la IBM, y SMP, del Instituto de Tecnología de California. Tales sistemas representan los máximos logros alcanzados hasta el presente en sistemas de cálculo algebraico computarizado.

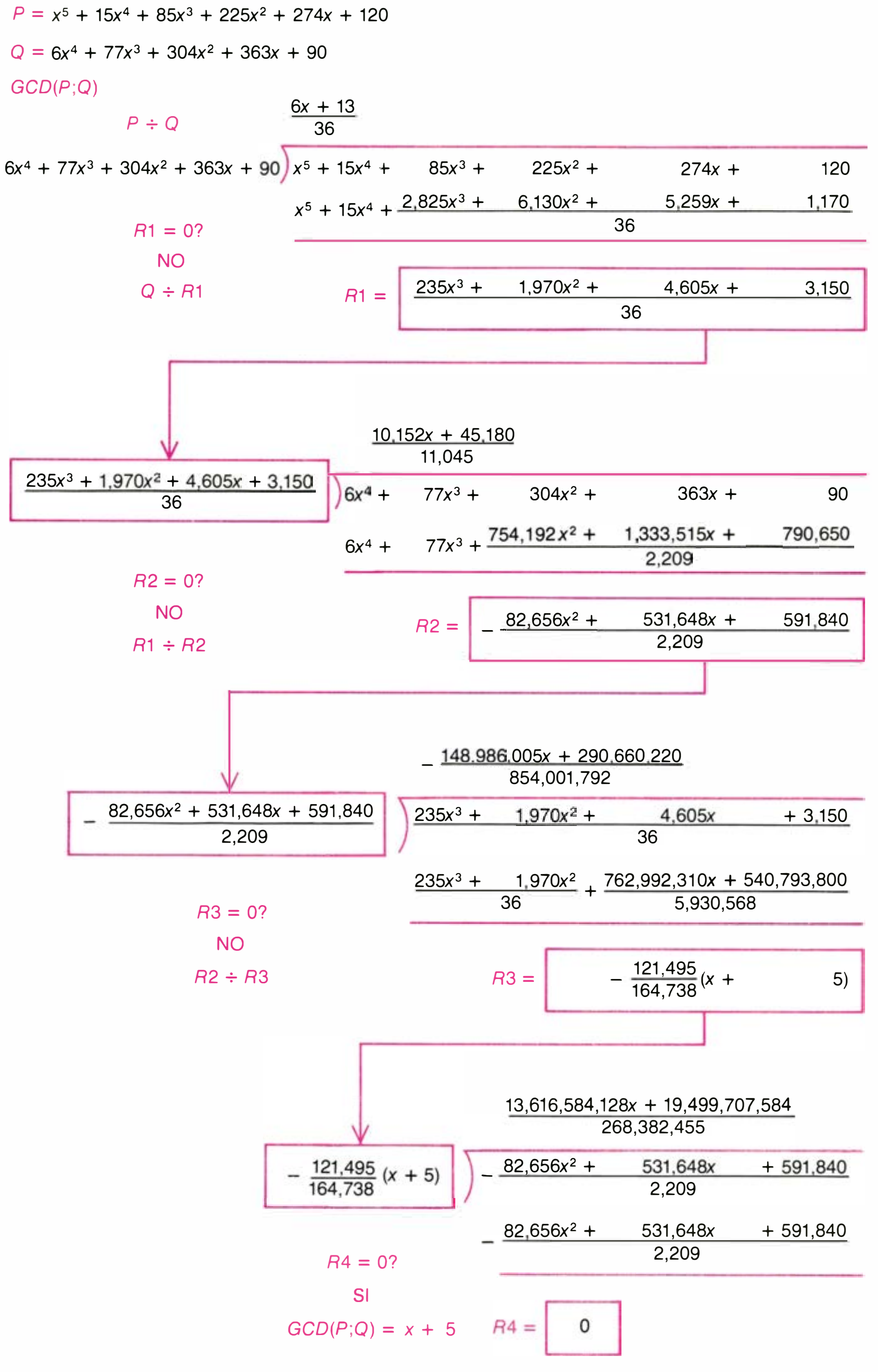
Están empezando a aparecer sistemas de cómputo algebraico diseñados para operar con microordenadores. Aunque más lentos y no tan extensos como los sistemas generales diseñados para macroordenadores, son capaces de realizar operaciones más complejas –y con mejor precisión– que la mayoría de los matemáticos. Entre tales sistemas, el más perfecto, más difundido y más accesible es el llamado MUMATH, preparado por la Soft Warehouse de Honolulú. El sistema MUMATH proporciona algunos de los recursos disponibles en sistemas generales; empero, el tamaño, capacidad de memoria y velocidad de los microordenadores no permiten atacar con MUMATH problemas verdaderamente complejos. A pesar de todo, tales sistemas (o descendientes suyos de mayor potencia) conseguirán seguramente abrirse paso hasta los microordenadores "de uso personal" y quién sabe si hacia finales de nuestro decenio no estarán ya disponibles en las pequeñas calculadoras de bolsillo.

Cada vez es más fuerte la tendencia a redactar los sistemas de operato-



de dos números enteros. (El máximo común divisor de dos números es el mayor de los enteros que dividen exactamente, sin resto, a ambos.) Hace al menos 2200 años que se descubrió el algoritmo; con la excepción de las cuatro operaciones elementales, se trata del más antiguo de los algoritmos conocidos. Las flechas rojas del diagrama de flujo indican los estados que va recorriendo el algoritmo al aplicarlo al ejemplo numérico de la parte inferior. El algoritmo de Euclides no sólo es aplicable a pares de enteros; con ligeras modificaciones, el procedimiento determina el máximo común divisor de dos polinomios.

ria algebraica computarizada en lenguajes de programación de alto nivel, que remedan las expresiones algebraicas ordinarias. Además, estos sistemas se han diseñado (por lo general) para interactuar con el usuario, sosteniendo con él una especie de conversación. Es frecuente que los noveles aprendan a resolver muchos tipos de problemas no triviales tras algunos minutos de práctica, problemas que rebasan con mucho la capacidad del aprendiz para resolverlos manualmente. En casi todos los sistemas generales, el usuario introduce sus órdenes en la máquina por medio de un teclado, sirviéndose de un lenguaje parecido a los conocidos FORTRAN o ALGOL.



APLICADO A POLINOMIOS, el algoritmo de Euclides determina el máximo común divisor (GCD, en sus iniciales inglesas) de ambos por medio de división reiterada, de la misma forma que se calcula el máximo común divisor de dos enteros. El algoritmo puede originar enormes resultados numéricos intermedios, que no es lícito redondear ni truncar. Por este motivo, las operaciones aritméticas de un sistema de álgebra por ordenador deben tener precisión ilimitada. Por lo común, el espacio de memoria de ordenador que se reserva para alojar un número es siempre el mismo, lo que no sucede en los sistemas de álgebra computarizada. Esta característica de los sistemas de álgebra por ordenador atrae a los matemáticos que estudian las propiedades de los números grandes. Las flechas y expresiones señaladas en rojo se corresponden con las etapas del algoritmo representado en el diagrama de las dos páginas precedentes.

Si la instrucción contiene fracciones, exponentes, u otros símbolos que comúnmente ocupan más de una línea, el ordenador presenta a través de su terminal una versión reformada, acorde con el uso corriente. El usuario puede entonces ordenar que se realicen ciertas operaciones sobre la expresión, como sumarla consigo misma, elevarla a potencia, derivarla o integrarla. El ordenador ejecuta la operación, simplifica la expresión resultante y la imprime o presenta en un terminal de vídeo. Cuando la operación presenta alguna indeterminación (pongamos por caso, si debe extraerse la raíz positiva o la negativa) el ordenador interroga al usuario antes de proseguir.

Se ha aplicado el álgebra computarizada a múltiples disciplinas y problemas: acústica, geometría algebraica, economía, mecánica de fluidos, mecánica estructural y teoría de números; ha probado su utilidad en el diseño de hélices, cascos de barcos, aspas de helicóptero, microscopios electrónicos e integración de circuitos a gran escala. Describiremos aquí tres aplicaciones más de los sistemas de operatoria algebraica por ordenador.

Se está investigando activamente la física de plasmas, en relación con la puesta a punto de fuentes de energía basadas en la fusión nuclear. En tales investigaciones es preciso combinar técnicas algebraicas y numéricas. En efecto, las expresiones que describen el comportamiento de plasmas en el seno de campos magnéticos son de enorme complejidad, lo que dificulta aún más la comprensión de los fenómenos físicos. Lo que se ha echado a faltar es un procedimiento capaz de obtener para estos fenómenos expresiones aproximadas útiles, más que complejos desarrollos analíticos exactos. Para vencer tal dificultad, el Grupo de Teoría de Plasmas del MIT ha ideado técnicas de cálculo algebraico por ordenador capaces de descomponer los términos del cómputo de acuerdo con los procesos físicos que los generan. Valiéndose después de métodos numéricos se consigue eliminar de la serie de términos que describen cada proceso aquellos considerados insignificantes. De esta forma se obtiene una expresión matemática que describe las propiedades dominantes, y que, al mismo tiempo, es lo bastante sencilla como para dejar ver el significado físico de cada uno de sus términos.

Para deducir predicciones, comprobables experimentalmente, a partir de las teorías de interacción de partículas

elementales, ha sido necesario calcular expresiones que bien pueden contener millares de integrales, muchas de las cuales encierran insólitas y engañosas propiedades matemáticas. El cálculo numérico de estas integrales padece de los dos defectos que ya hemos mencionado: ser mucho menos útiles que los resultados algebraicos, donde puede quedar claro cómo contribuyen a modificar la interacción los diversos fenómenos físicos y, por otra parte, introducir errores de redondeo y truncamiento. Conforme las técnicas experimentales se hacen más refinadas, las incertidumbres provocadas por las inexactitudes del cálculo pueden ser tan grandes que hagan indistinguibles las predicciones correspondientes a teorías distintas. Es de esperar que con la aplicación de métodos de operatoria computarizada puedan generarse predicciones teóricas de la precisión necesaria.

En el estudio de teorías alternativas a la relatividad general ha llegado a popularizarse como criterio el llamado teorema de Birkhoff. En 1923, el matemático George David Birkhoff, de la Universidad de Harvard, demostró que la teoría de relatividad general excluye la posibilidad de propagación de pulsos gravitatorios a través del espacio; tales brotes pudieran concebiblemente generarse por pulsaciones de sentido radial de la materia en ciertas estrellas. Puesto que no han podido detectarse ningún tipo de pulsos gravitatorios, y debido también a que tales brotes son imposibles en las teorías de Newton y Einstein, se ha dado en requerir que toda nueva teoría excluya también la posibilidad de dichos pulsos.

Los cálculos necesarios para determinar si una teoría gravitatoria verifica o no el teorema de Birkhoff resultan prolijamente largos, pero gracias a los sistemas de operatoria computarizada la cuestión puede dirimirse sin gran dificultad. Recientemente, por ejemplo, se ha comprobado por este procedimiento que una teoría gravitatoria creada por C. N. Yang, de la Universidad estatal de Nueva York en Stony Brook, violaba el principio de Birkhoff. El nuevo cálculo sirvió para corregir otro realizado a mano, que fue admitido como correcto y citado como tal durante no pocos años.

234 THÉORIE DU MOUVEMENT DE LA LUNE.

la fonction R ne contient plus aucun terme périodique; elle se trouve donc réduite à son terme non périodique seul, terme qui, en tenant compte des parties fournies par les opérations 129, 260, 349 et 415, a pour valeur

$$R = \frac{\mu}{2a}$$

$$+ m'\frac{a^2}{a'^3}\Big\{ \frac{1}{4} - \frac{3}{2}\gamma^2 + \frac{3}{8}e^2 + \frac{3}{8}e'^2 + \frac{3}{2}\gamma^4 - \frac{9}{4}\gamma^2e^2 - \frac{9}{4}\gamma^2e'^2 + \frac{9}{16}e^2e'^2 + \frac{15}{32}e'^4 - \frac{33}{2}\gamma^4e^2$$

$$+ \frac{9}{4}\gamma^4e'^2 + \frac{75}{16}\gamma^2e^4 - \frac{27}{8}\gamma^2e^2e'^2 - \frac{45}{16}\gamma^2e'^4 + \frac{45}{64}e^2e'^4$$

$$+ \Big(\frac{9}{16}\gamma^2 + \frac{225}{64}e^2 - \frac{27}{16}\gamma^4 - \frac{387}{32}\gamma^2e^2 + \boxed{\frac{23}{16}\gamma^2e'^2} - \frac{225}{128}e^4 + \frac{825}{64}e^2e'^2 + \frac{9}{8}\gamma^6$$

$$+ \frac{3897}{64}\gamma^4e^2 - \frac{99}{16}\gamma^4e'^2 - \frac{1431}{256}\gamma^2e^4 - \frac{1419}{32}\gamma^2e^2e'^2 - \frac{225}{512}e^6 - \frac{825}{128}e^4e'^2\Big)\frac{n'}{n}$$

$$- \Big(\frac{31}{32} - \frac{33}{8}\gamma^2 - \frac{971}{32}e^2 + \frac{465}{64}e'^2 + \frac{273}{64}\gamma^4 + \frac{5709}{64}\gamma^2e^2 - \frac{117}{4}\gamma^2e'^2 + \frac{4989}{256}e^4$$

$$- \frac{1905}{8}e^2e'^2 + \frac{3255}{128}e'^4\Big)\frac{n'^2}{n^2}$$

$$- \Big(\frac{255}{32} - \frac{31515}{1024}\gamma^2 - \frac{551115}{4096}e^2 + \frac{6885}{64}e'^2 + \frac{20511}{512}\gamma^4 + \frac{927831}{2048}\gamma^2e^2$$

$$- \frac{218115}{512}\gamma^2e'^2 + \frac{1622985}{16384}e^4 - \frac{4069635}{2048}e^2e'^2\Big)\frac{n'^3}{n^3}$$

$$- \Big(\frac{5515}{192} - \frac{296779}{3072}\gamma^2 - \frac{6380965}{12288}e^2 + \frac{16285}{24}e'^2\Big)\frac{n'^4}{n^4}$$

$$- \Big(\frac{28841}{288} - \frac{1138183307}{294912}\gamma^2 - \frac{1681901051}{1179648}e^2 + \frac{1393609}{384}e'^2\Big)\frac{n'^5}{n^5}$$

$$- \frac{9814775}{36864}\frac{n'^6}{n^6} - \frac{428268199}{663552}\frac{n'^7}{n^7}$$

$$+ \Big[\frac{9}{64} - \frac{45}{16}\gamma^2 + \frac{45}{64}e^2 + \frac{15}{128}e'^2$$

$$+ \Big(\frac{225}{512} - \frac{1935}{256}\gamma^2 + \frac{7425}{1024}e^2 + \frac{225}{64}e'^2\Big)\frac{n'}{n} + \frac{869}{512}\frac{n'^2}{n^2} - \frac{10391}{8192}\frac{n'^3}{n^3}\Big]\frac{a^2}{a'^2}\Big\}.$$

UN GIGANTESCO CALCULO, realizado manualmente y terminado en 1867, se comprobó en 1970 con la ayuda de un sistema de álgebra por ordenador. El cálculo primitivo lo efectuó Charles Delaunay, y se publicó en dos volúmenes. Vemos reproducida aquí una página del segundo volumen. Para realizar el cálculo, Delaunay precisó de 10 años, y de otros 10 para repasarlo; con él se proponía dar la posición de la Luna en función del tiempo con precisión jamás lograda hasta entonces. Al repasar el cálculo, André Deprit, Jacques Henrard y Arnold Rom, de los Boeing Scientific Research Laboratories, de Seattle, quedaron al descubierto tres errores. El principal, causante de los otros dos, aparece aquí recuadrado en rojo; la expresión correcta es 33/16 $\gamma^2 e'^2$. El ordenador necesitó 20 horas de funcionamiento para rehacer los cálculos. Aunque un sistema de álgebra computarizada ha sido el responsable de mostrar los errores de Delaunay, las tornas se han vuelto: dada su enorme precisión, el cálculo de Delaunay se utiliza para comprobar la exactitud de nuevos sistemas de álgebra por ordenador antes de ponerlos en servicio.

Conviene tener en cuenta que la utilidad de los sistemas de álgebra computarizada no sólo se manifiesta en cálculos extensos. Con frecuencia se presentan problemas relativamente sencillos, donde la manipulación algebraica puede facilitar la comprensión o, tal vez, hacer más eficiente el cálculo numérico de sus soluciones. Aplicando juiciosamente métodos algebraicos, no pocas veces se consigue dividir por 100 el tiempo de ordenador destinado a procesar un problema numérico. Además, los programas de álgebra computarizada pueden reemplazar con ventaja toda una biblioteca de obras de consulta y referencia. Los programas de "conocimientos residentes", al disponer de métodos de integración, de métodos para resolver ecuaciones diferenciales y otros semejantes pueden dejar tan obsoletas a estas obras de referencia como las modernas calculadoras de mano han dejado ya a las tablas de logaritmos. Conforme vayan perfeccionándose los ordenadores y reduciéndose sus costos, los sistemas de álgebra computarizada serán cada vez más accesibles a la enseñanza, estudio, investigación y cualesquiera otras insospechadas aplicaciones en todos los hogares y oficinas donde haya personas interesadas.

El principio antrópico

Ciertas condiciones, así la temperatura, fueron favorables a la aparición de la vida en la Tierra. El principio antrópico razona al revés: la presencia de vida puede "explicar" tales condiciones

George Gale

La Tierra es un lugar excepcionalmente hospitalario para la humanidad, con abundancia de agua y una temperatura cuya media se mantiene, por fortuna, dentro de la reducida escala de grados a los que el agua es líquida. Atendido el origen evolucionista de la vida, estos hechos no son muy de extrañar: si la Tierra fuese fría y seca como Marte, o si tuviese una atmósfera gaseosa, cáustica y tórrida, como la de Venus, no habrían evolucionado en ella seres inteligentes que pudieran comentar lo inhóspito de su entorno. Lo que sí parece, en cambio, francamente raro es sostener que la presencia de vida sobre la Tierra pueda "explicar" por qué el planeta tiene una temperatura entre el punto de congelación y el de ebullición del agua. La práctica usual ha sido sostener la tesis opuesta, o sea, que la vida evolucionó aquí porque las circunstancias fueron propicias a su existencia.

Aunque el razonamiento quizá parezca al principio chocante, la idea de que la mera presencia de vida inteligente puede tener algún poder explicativo ha sido introducida recientemente en cosmología, donde se trata de entender la historia no de un solo planeta sino del universo entero. Es fácil imaginar un universo totalmente diferente del que tenemos observado. Por ejemplo, el cambiar los valores de ciertas constantes físicas podría dar origen a un universo donde nunca se hubiesen formado los elementos químicos más pesados que el helio o donde todas las estrellas fuesen grandes, calientes y de corta duración. En la mayoría de tales reconstrucciones imaginarias del universo es improbable que apareciera nunca una forma de vida inteligente. El hecho de que el universo real albergue a observadores inteligentes pone, pues, algunas restricciones a la diversidad de comienzos posibles del universo y a las leyes físicas que podrían haber regido su desarrollo. En otras palabras, el universo tiene las propiedades que hoy observamos porque, si las de su más temprana infancia hubiesen sido muy diferentes, no estaríamos aquí nosotros contemplándolo. El principio subyacente a este método de análisis cosmológico ha sido denominado principio antrópico, del griego *anthropos*, hombre.

La manera de razonar basada en el principio antrópico es completamente distinta de la deductiva, que tanto tiempo ha caracterizado a gran parte del pensamiento científico. Una teoría deductiva empieza por especificar las condiciones iniciales de un sistema físico y las leyes de la naturaleza que se le aplican; la teoría predice a continuación el subsiguiente estado del sistema. Por ejemplo, podría alguien deducir las condiciones que se dan al presente sobre la Tierra precisando primero el tamaño inicial, la masa y la composición química de la nebulosa a partir de la cual se condensó el sistema solar, rastreando después la evolución del Sol y de los planetas bajo el influjo de leyes físicas que describen fuerzas gravitacionales, reacciones nucleares y así sucesivamente. El principio antrópico ha sido invocado en cosmología precisamente porque aquí no es fácil emplear el método deductivo. Se desconocen las condiciones iniciales del universo y son también inciertas las leyes físicas que operaron en los albores de su historia; puede, incluso, que las leyes dependieran de las condiciones iniciales. A decir verdad, tal vez la única norma que quepa imponer a una teoría que intente reconstruir las condiciones iniciales del universo y las correspondientes leyes de la naturaleza sea el requisito de que tales condiciones y leyes den origen a un universo habitado.

Por lo menos el principio antrópico sugiere conexiones entre la existencia del hombre y aspectos de la física que podría haberse pensado que no tenían mucho que ver con la biología. En su forma más fuerte, el principio vendría a revelarnos que el universo en que vivimos es el único concebible donde pueda existir vida inteligente. Hay que decir, sin embargo, que no todos los cosmólogos y filósofos de la ciencia están de acuerdo en cuanto a la utilidad del principio antrópico, ni siquiera en cuanto a su legitimidad. Describiré aquí algunas de las maneras como ha sido aplicado y dejaré que el lector juzgue acerca de su validez.

El concepto que subyace a mucha de la cosmología moderna se llama principio copernicano. Sus orígenes se remontan a la afirmación, hecha por Nicolás Copérnico en 1543, de que la Tierra no es el centro del universo. Pero la forma como modernamente se ha difundido el principio no fue explicitada hasta 1948, año en que lo formuló Hermann Bondi, de la Universidad de Cambridge. Sostiene que la posición de los observadores humanos en el universo no es en modo alguno privilegiada o distinta de otras posiciones; por lo que, en cosmología, las observaciones son válidas no sólo para la Tierra o para el sistema solar, sino también para remotas regiones del universo. El principio copernicano o algún supuesto similar es metodológicamente necesario en cosmología; sin él los resultados de la investigación cosmológica podrían tacharse de deformaciones derivadas de las características peculiares a la fracción del universo donde habitan los observadores humanos. Ni que decir tiene que, según reconociera el mismo Bondi, la utilidad del principio copernicano no es ninguna garantía de su verdad.

Una generalización del principio copernicano ha venido a ser conocida como el principio cosmológico. Dice éste que no sólo es la posición del sistema solar la que no tiene estatuto alguno de privilegio, sino que tampoco lo tiene ninguna otra posición en cualquier parte que sea del universo. Esta idea lleva

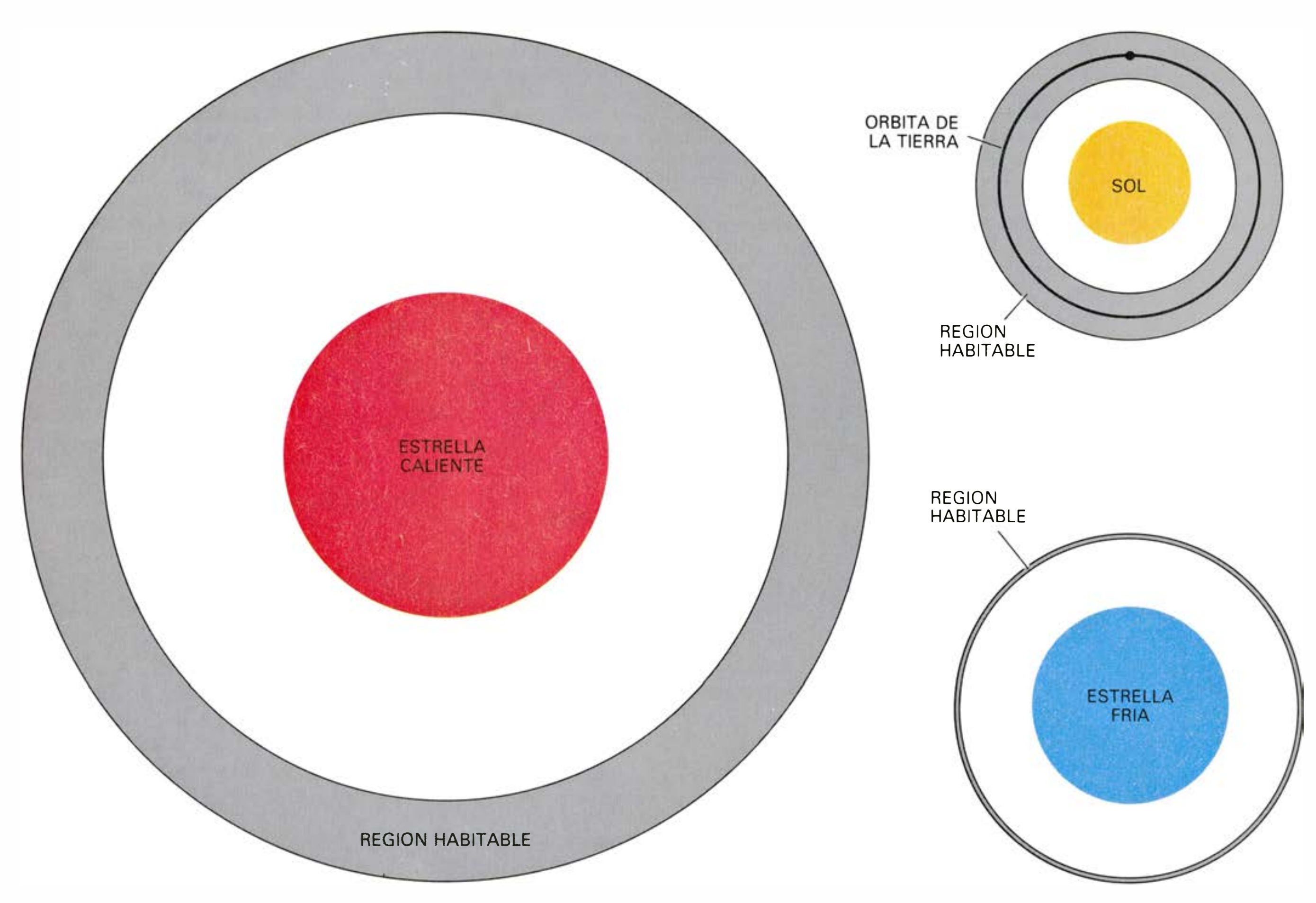
implícito el supuesto de que, a gran escala, la estructura del universo es uniforme: aparte de irregularidades locales, tales como las galaxias, todas las regiones del universo son exactamente iguales. Una estructura homogénea es atractiva (a falta de pruebas en contrario), porque es la más simple estructura posible. Sobre este supuesto metodológico, la Tierra ocupa una típica posición en el espacio.

Las pruebas en pro del principio cosmológico dimanan de la reproducibilidad de la mayoría de los experimentos científicos. Aun cuando un experimento, una medición de la velocidad de la luz por ejemplo, se efectúe repetidamente en el mismo laboratorio, se estará efectuando, con todo, en diferentes tiempos y en diferentes puntos del espacio (porque en el *interin* se ha ido moviendo la Tierra). En cuanto los resultados son los mismos, la posición de la Tierra no afecta de hecho al experimento. Sin embargo, tal evidencia no es muy convincente, pues en cosmología las conclusiones atañen a regiones del espacio-tiempo mucho mayores que las recorridas por nuestro planeta.

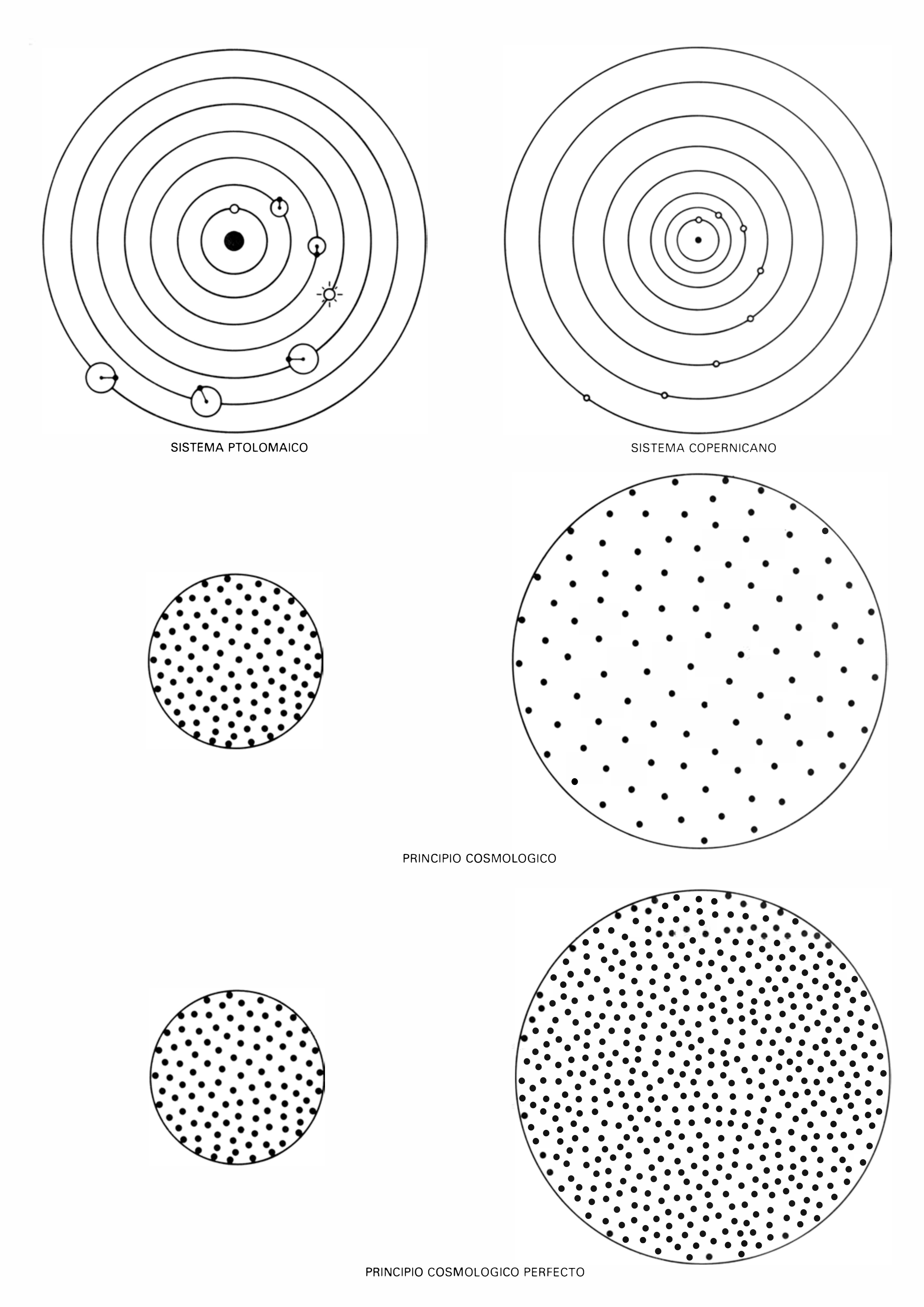
Bondi y Thomas Gold, de la Universidad de Cornell, propusieron una hipótesis aún más general denominada el principio cosmológico perfecto. Afirma este que, prescindiendo de irregularidades locales, el universo es uniforme en el espacio y en el tiempo, de suerte que un observador vería la misma estructura a gran escala desde cualquier lugar y en cualquier época. El principio cosmológico perfecto es la base del modelo estacionario, que en su formulación original postulaba un universo completamente uniforme en todo el espacio y el tiempo. Para adaptarse a la evidencia de que el universo se halla en expansión, el modelo supone que continuamente se está creando materia. Hace ya bastante que se abandonó el modelo estado-estacionario, y, con él, lo ha sido también el principio cosmológico perfecto. La causa del abandono fue la detección, en 1965, de la radiación universal de fondo de microondas. La radiación de fondo de microondas se interpreta como el remanente de un estadio del universo primitivo, cuando este era mucho más cálido y más denso que ahora.

Aunque la radiación de fondo de microondas excluye la uniformidad temporal del universo, proporciona en cambio la prueba más fuerte a favor de la uniformidad espacial a gran escala. La radiación observada es isótropa –esto es, viene de todas las direcciones con igual intensidad– a una precisión superior al 1 por 1000. Así que las propiedades de la radiación de fondo confirman el principio cosmológico, pero no el principio cosmológico perfecto.

La observada expansión del universo



SISTEMAS ESTELARES HABITABLES son sólo aquellos en los que una estrella calienta una región del espacio donde la temperatura no pasa de la estrecha franja en la que el agua se mantiene líquida. En el diagrama de la izquierda está representada la más caliente estrella cuyos planetas podrían sustentar vida inteligente; en el diagrama de la derecha, arriba, el Sol, y, abajo, la estrella más fría cuyos planetas serían capaces de sustentar vida inteligente. (En cada diagrama, se han dibujado a escala diferente la estrella y el sistema planetario.) Los límites superior e inferior del tamaño de una estrella capaz de sostener vida dependen de las distintas propiedades de los sistemas estelares. Una estrella mayor y más brillante podría proporcionar una región habitable, pero la estrella permanecería en un estadio estable de su evolución un tiempo insuficiente para que la vida evolucionara. Una estrella más pequeña y más oscura podría también proporcionar una región habitable, pero el planeta estaría demasiado próximo a la estrella. A resultas de las interacciones de fuerzas atractivas el planeta dejaría de rotar y se produciría una extrema diferencia de temperaturas entre su cara iluminada y su cara en sombra, hasta que, por último, la atmósfera del planeta sería hirviente en el lado luminoso y gélida en el oscuro. La vida llegó a formarse en la Tierra gracias a la circunstancia de que nuestro planeta ocupa una región habitable. El principio antrópico, desarrollado aquí, respalda la tesis opuesta: la presencia de vida en la Tierra explica por qué la temperatura de nuestro planeta se mantiene dentro de la estrecha franja en la que el agua es líquida.


SISTEMA PTOLOMAICO
SISTEMA COPERNICANO
PRINCIPIO COSMOLOGICO
PRINCIPIO COSMOLOGICO PERFECTO

es también compatible con el principio cosmológico. La expansión no tiene ningún centro: un observador situado en cualquier galaxia vería alejarse de él en todas direcciones a las demás galaxias distantes. Cuanto más lejana se halla una galaxia, mayor es su velocidad de recesión. Para las galaxias que están a la misma distancia son iguales las velocidades de recesión, dentro de un margen de exactitud superior al 1 por 1000.

La radiación de fondo de microondas, la recesión de las galaxias distantes y el principio cosmológico se conjuntan todos en el modelo de la gran explosión, que toma como origen del universo un punto sin dimensiones y de densidad aparentemente infinita. Tal origen se sugiere al imaginar que se inviertan las actuales velocidades de recesión; extrapolando hacia atrás se llegaría a que todas las galaxias se encerraran en un punto. Como el universo se ha venido expandiendo desde el instante de la gran explosión, puede calcularse su edad por las características de dicha expansión. Si las velocidades de recesión no hubiesen cambiado a lo largo del tiempo, la edad del universo sería igual a la distancia entre dos galaxias cualesquiera dividida por la velocidad a que se van alejando una de otra. A esta edad hipotética se le llama tiempo de Hubble, en honor de Edwin P. Hubble, quien, en 1923, descubrió la relación existente entre la distancia y la velocidad de recesión.

De hecho, es probable que las velocidades de recesión hayan cambiado, aunque no está claro en qué cuantía. Una causa de tal cambio es la mutua atracción gravitatoria de todas las galaxias, que tiende a decelerar la expansión del universo. El cambio de la velocidad de expansión en el período transcurrido desde la gran expansión determina si el universo es "abierto" o "cerrado". Un universo abierto seguirá expandiéndose a perpetuidad. Un universo cerrado cesará de expandirse en algún momento, empezará a contraerse y, por último, llegará al colapso parándose del todo en la "gran implosión". Si el universo es abierto, la actual velocidad de expansión sugiere que su edad es de unos 20.000 millones de años. Si es cerrado, la velocidad de expansión parece indicar que su edad ronda los 13.000 millones de años.

El principio antrópico fue introducido por Robert H. Dicke, de la Universidad de Princeton, en 1961; lo propuso al analizar el trabajo realizado por P. A. M. Dirac unos 30 años antes. Dirac había llamado la atención sobre ciertas curiosas relaciones numéricas entre números sin dimensiones que tienen un importante papel en física y en astrofísica. Número sin dimensiones es aquel que no está asociado a ninguna unidad de medida, con lo que su valor es el mismo en cualquier sistema de medición. Dirac no consideró el valor exacto de los números, sino sólo su orden de magnitud: la potencia de 10 que más aproximadamente expresa el valor. Encontró varios ejemplos en los que el orden de magnitud es una potencia integral del enorme número 10^{40}.

Tres números que figuraban prominentemente en el trabajo de Dirac son medidas de fuerza, tiempo y masa. La primera cantidad es una forma adimensional de la constante de interacción gravitatoria, que es una medida de la intensidad de la fuerza gravitacional; tiene un valor aproximado de 10^{-40}. El segundo número sin dimensiones es la edad del universo expresada en unidades atómicas: Dirac lo definió como la razón de la edad de Hubble al tiempo que tarda la luz en recorrer una distancia igual al radio de un protón. El valor aproximado de esta razón es 10^{40}. (Como Dirac atendía solamente al orden de magnitud, los cálculos de la edad de Hubble y los de la otra edad dan en gran parte el mismo resultado.) La tercera cantidad adimensional es el número de partículas con masa (protones y neutrones) que hay en la región visible del universo; este número se calcula que es aproximadamente de 10^{80}.

Dirac señaló tres relaciones ordomagnitudinales entre estas cantidades. Primera, que la constante de interacción gravitacional es el recíproco de la edad del universo en unidades atómicas. Segunda, que el número de las partículas con masa es el cuadrado de la edad del universo en unidades atómicas. Tercera, que la constante de interacción gravitacional es el recíproco de la raíz cuadrada del número de partículas con masa. Dirac pensó que tales relaciones numéricas eran tan llamativas que no se las podía despachar como meras coincidencias casuales. Propuso, pues, la hipótesis de que fueran el resultado de alguna conexión causal desconocida.

Una objeción posible contra estas ideas es la de que, obviamente, la edad del universo aumenta con el tiempo. Por lo que las relaciones entre aquellos números tendrán que ir cambiando sin cesar, y es una extraordinaria coincidencia el que sus valores se determinen precisamente cuando se dé el caso de que correspondan entre sí. Dirac se anticipó a esta crítica proponiendo la hipótesis de que la constante de interacción gravitacional y el número de partículas con masa cambien también con el tiempo, de tal modo que las relaciones ordomagnitudinales sigan siendo válidas a lo largo de toda la historia del universo. Para que persistan las correspondencias habrá de irse debilitando la gravedad en proporción inversa al tiempo, mientras que el número de partículas deberá aumentar en proporción directa al cuadrado del tiempo.

El análisis de Dirac fue recibido, en general, con poco entusiasmo; pero Dicke se lo tomó en serio. Y propuso que una conexión causal entre la constante de interacción gravitacional y el número de partículas con masa podría tener su fundamento en un principio enunciado primeramente por Ernst Mach. Este había avanzado que la masa inercial de una partícula venía determinada por su interacción gravitatoria con la materia distante. (La opinión admitida era que la masa inercial constituía una propiedad de la partícula enteramente independiente de su entorno.) De acuerdo con el principio de Mach, la debilidad de la gravitación está relacionada con la enorme cantidad de materia distante en el universo. Si se acepta el principio, nada tiene de extraño que haya alguna relación numérica en-

MODELOS DE UNIVERSO prevalecientes en distintas épocas de los últimos 2000 años; estos diagramas los representan en orden de creciente simetría. En el sistema ptolemaico, codificado en el s. II d.J.C., la Tierra ocupa una posición fija en el centro del universo y todos los demás cuerpos celestes giran en círculo a su alrededor. En 1543 Nicolás Copérnico sugirió que el centro del universo no era la Tierra, sino el Sol. En el siglo XX se ha descubierto que el universo se expande uniformemente en todas direcciones, y por consiguiente se ha abandonado la idea de que el universo tenga un centro. El principio cosmológico afirma que las características a gran escala del universo le parecerían las mismas al observador que en cualquier galaxia mirase en cualquier dirección. El diagrama esquemático del principio cosmológico muestra la expansión del universo en dos momentos; cada punto representa una galaxia. Elíjase la galaxia que se elija, todas las demás se van apartando de ella; a medida que el universo se expande, cambia la distancia entre las galaxias, pero no la distribución geométrica de las mismas. El principio cosmológico perfecto presenta un universo que es aún más simétrico: sus rasgos a gran escala son los mismos no sólo desde cualquier punto del espacio, sino también desde cualquier instante del tiempo. En la versión del principio cosmológico perfecto aquí esquematizada el universo se expande; no obstante, se van formando nuevas galaxias precisamente a la velocidad necesaria para que se mantenga una densidad constante. El principio cosmológico perfecto es hoy en gran parte desechado. El principio antrópico, no diagramado aquí, afirma que la Tierra está en una posición privilegiada porque hay en nuestro planeta vida inteligente.

tre la constante de interacción gravitacional y el número de partículas con masa, que es una medida de la cantidad de materia que hay en el universo.

No está claro, empero, por qué la constante gravitacional y el número de partículas tengan que relacionarse con la edad de Hubble. Por el contrario, si el principio de Mach es totalmente válido, deberá serlo en todas las eras de la historia del universo, mientras que las relaciones ordomagnitudinales se observarían sólo en la era actual. Nuevamente parece que el hombre ha aparecido en un momento privilegiado y, por tanto, poco probable.

La respuesta de Dicke a esta objeción fue que el valor de la edad de Hubble está muy forzado por las condiciones necesarias para la existencia del hombre. Una condición esencial es que el universo sea lo bastante antiguo como para haber dado tiempo a la formación de elementos más pesados que el hidrógeno, pues "ya se sabe que ha de haber carbón para que haya físicos". Los elementos pesados se forman en el interior de las estrellas y salen liberados cuando una estrella estalla en fase de supernova. Por consiguiente, la edad de Hubble de un universo habitado no puede ser menor que la de la estrella de más corta vida. Por otro lado, si la edad de Hubble fuese mucho mayor que la de una estrella típica, la mayoría de las estrellas cuyos planetas pudiesen sustentar vida se habrían apagado ya. Por lo tanto –concluía Dicke– la edad de Hubble es aproximadamente igual a lo que dura la vida de una estrella típica.

Lo novedoso del argumento de Dicke merece detenido análisis. Dada la existencia del hombre –argüía– pudiera ser que la edad de Hubble no tuviese un valor muy diferente del que actualmente tiene. De ahí que las relaciones numéricas de Dirac se apliquen no a cualquier universo evolutivo posible (donde la edad de Hubble podría presumiblemente tomar uno cualquiera de entre muchos valores), sino sólo al universo que observamos hoy.

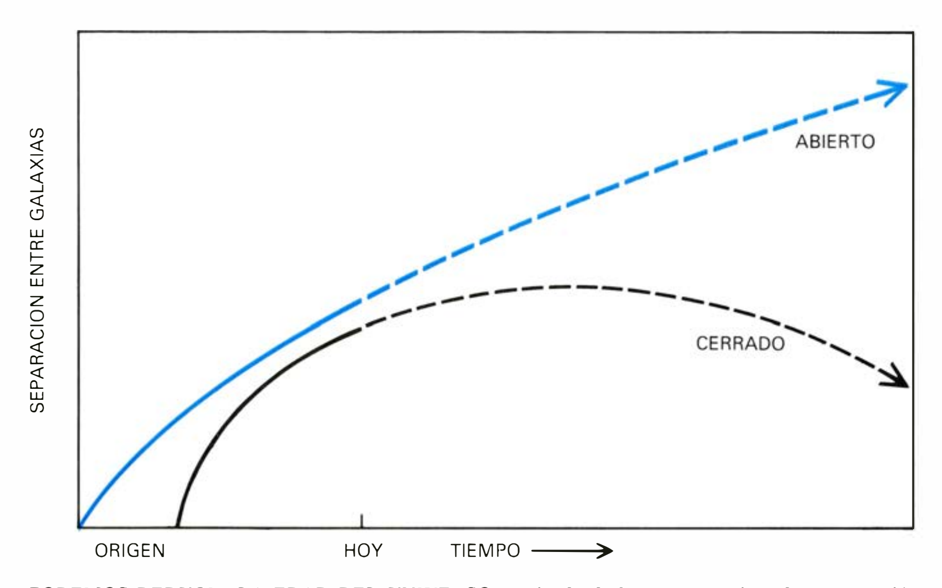
Uno de los rasgos más atractivos del análisis de Dicke es su clara demostración de que el valor de la edad de Hubble no es arbitrario. El reducir la arbitrariedad de las explicaciones constituye una vieja aspiración de las ciencias, y así, en este aspecto, el trabajo de Dicke no es insólito; lo que hace que sea distinto es el método o la lógica de la argumentación. Generalmente, la arbitrariedad ha sido eliminada demostrando que un fenómeno puede predecirse o que una teoría puede deducirse de alguna premisa más fundamental. La técnica de Dicke es por completo diferente.

La lógica deductiva o predictiva procede desde una suposición fundamental hasta un resultado derivado: el futuro es deducido del pasado. El flujo temporal del argumento de Dicke va en la dirección opuesta. El cita una condición presente (la existencia del hombre) como explicación de un fenómeno que tiene su base en el pasado (la edad del universo). Evidentemente, el resultado al que llega no puede interpretarse como una predicción, pues sería una predicción del pasado basada en el propio futuro de aquel pasado.

Los cosmólogos se han vuelto hacia el principio antrópico porque es difícil aplicar la lógica predictiva al universo primitivo. Una explicación deductiva en cosmología querría mostrar presumiblemente cómo rasgos que se observan en el universo, por ejemplo la distribución de la materia o el valor de la constante de interacción gravitacional, no son arbitrarios, sino que se siguen más bien de algún principio subyacente. Tal explicación es difícil darla, porque requiere conocer las condiciones iniciales del universo.

Un rasgo observado del universo que sigue necesitando explicación es el de su isotropía. C. B. Collins y Steven W. Hawking, de la Universidad de Cambridge, han encontrado que en los actuales modelos de universo sólo unas pocas series de condiciones iniciales, de entre las muchas condiciones posibles, podrían dar origen a la isotropía observada. Toda teoría en la que se deduzca o prediga la isotropía ha de comenzar postulando unas condiciones iniciales tan sumamente arbitrarias. Collins y Hawking estiman que este resultado es insatisfactorio, pues no ofrece ninguna razón que obligue a pensar que el universo ha seguido el curso que ahora sigue y no otro. Lo que se necesita es alguna constricción anterior que explique por qué las condiciones iniciales hubieron de estar entre aquellas pocas que condujeron a la isotropía; sólo que un previo imperativo sobre las condiciones iniciales del universo es casi inconcebible. De ahí que los investigadores hayan recurrido al principio antrópico, que limita la clase de las posibles condiciones iniciales, no mediante una constricción anterior, sino mediante una subsiguiente.

Brandon Carter, de Cambridge, ha efectuado una de las más influyentes aplicaciones del principio antrópico y de la lógica no predictiva. Carter comenzó por explorar la línea antrópica de investigación en "reacción contra el exagerado sometimiento al principio copernicano". Arguye Carter que, si bien Copérnico demostró que no debemos "dar gratuitamente por supuesto que ocupamos una posición central en el universo", no se sigue de ello que los

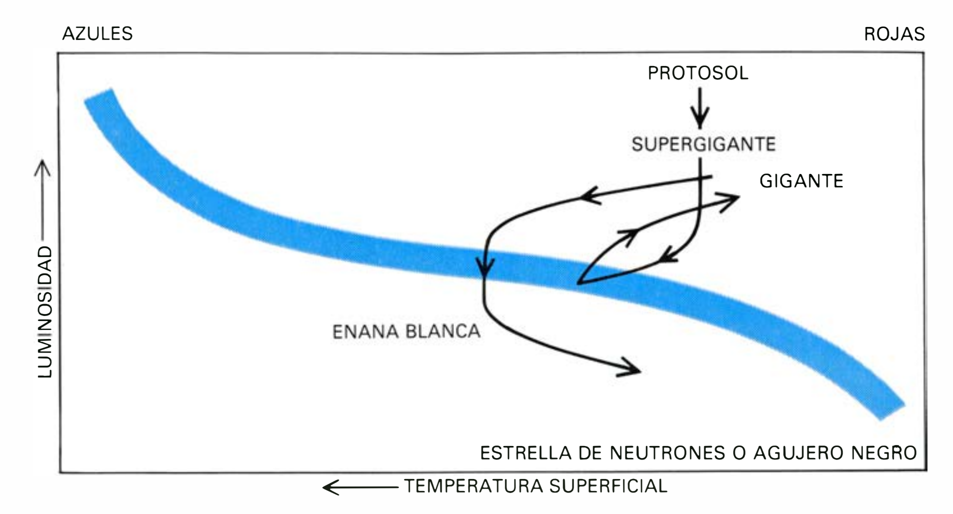


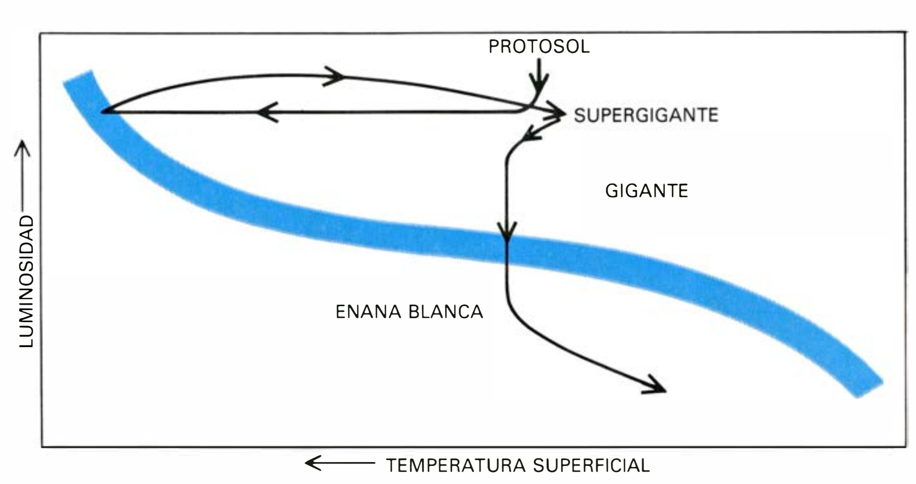
PODEMOS DEDUCIR LA EDAD DEL UNIVERSO partiendo de las características de su expansión actual. El esquema representa la separación entre las galaxias como una función del tiempo. El punto del "Hoy" corresponde a la actual velocidad de expansión. Pero sólo a partir de ésta no puede deducirse la edad del universo: probablemente la velocidad ha venido disminuyendo desde la gran explosión, a causa de la atracción gravitatoria de la materia en expansión. No se sabe si la expansión continuará por siempre (*curva coloreada*) o si el universo dejará en algún momento de expandirse, para ir entonces encogiéndose por su propia gravitación (*curva negra*). Ambas posibilidades son compatibles con las pruebas de que se dispone y con el modelo de la gran explosión. La expansión continua sugiere una edad de unos 20.000 millones de años; la expansión seguida de repliegue da una edad de alrededor de los 13.000 millones de años. La edad que se calcula suponiendo una velocidad de expansión constante es el tiempo de Hubble.

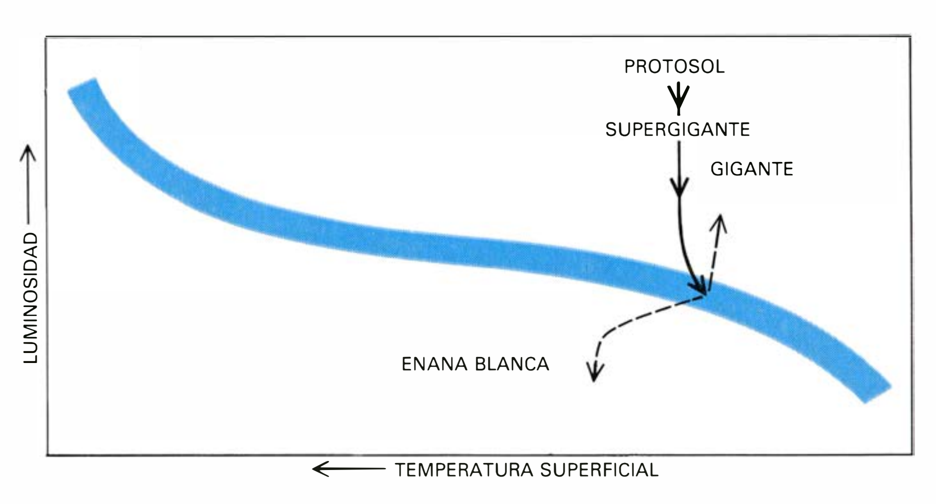
observadores humanos no puedan en modo alguno ser privilegiados. La posición del observador es necesariamente especial al menos en tanto en cuanto ciertas condiciones (de temperatura, ambiente químico y otras más) son prerrequisitos de su existencia. "Lo que cabe que esperemos observar", advierte Carter, "ha de hallarse limitado por las condiciones que son necesarias para que se dé nuestra presencia como observadores", y así, "aunque nuestra situación no sea necesariamente central, es inevitablemente privilegiada hasta cierto punto".

El tratamiento que hace Carter del principio antrópico representa una interesante conjunción de la física de lo muy grande y la de lo muy pequeño; para ilustrar la cosmología se fía de una curiosa explicación de la mecánica cuántica denominada interpretación pluricósmica (de muchos mundos). Fue ésta propuesta por Hugh Everett III, de Princeton, y desarrollada ulteriormente por Bryce S. De Witt y John Archibald Wheeler, de la Universidad de Texas en Austin. En la teoría cuántica las predicciones dan sólo la probabilidad de un evento y no una aserción determinista de si el evento ocurrirá o no. Por ejemplo, la trayectoria de una partícula elemental es descrita por una función de onda, expresión matemática cuya amplitud varía en espacio y en tiempo. La probabilidad de encontrar la partícula en un punto dado es igual al cuadrado de la amplitud de la función de onda en ese punto. Pero si se efectúa realmente una observación en el punto, a la partícula puede hallársela o puede no hallársela ahí. Un interés filosófico central en la mecánica cuántica es el de conciliar la interpretación probabilista de la función de onda con el resultado determinístico de las observaciones. Cuando la partícula es observada en una cierta posición, ¿tenía ella esta posición desde el principio, aun antes de que se hiciese la observación? Si la posición es determinística, no está claro cómo haya que interpretar los otros puntos del espacio a los que la función de onda asignaba una probabilidad no nula.

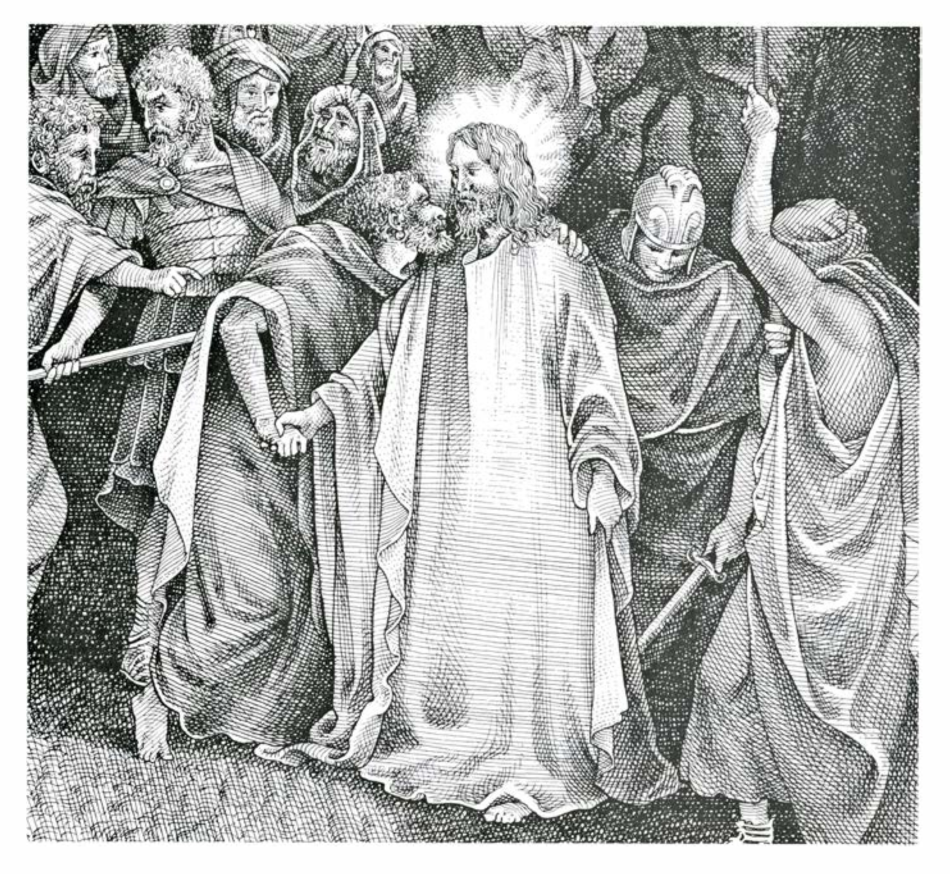
La interpretación pluricósmica de la mecánica cuántica asevera que no hay ninguna diferencia fundamental entre la posición observada de la partícula y los otros puntos a los que la función de onda asignaba una probabilidad nocero. La partícula existe en todos los puntos. Para que esto sea verdad es necesario, empero, suponer que hay una infinidad de mundos, en cada uno de los cuales la partícula tiene una posi-

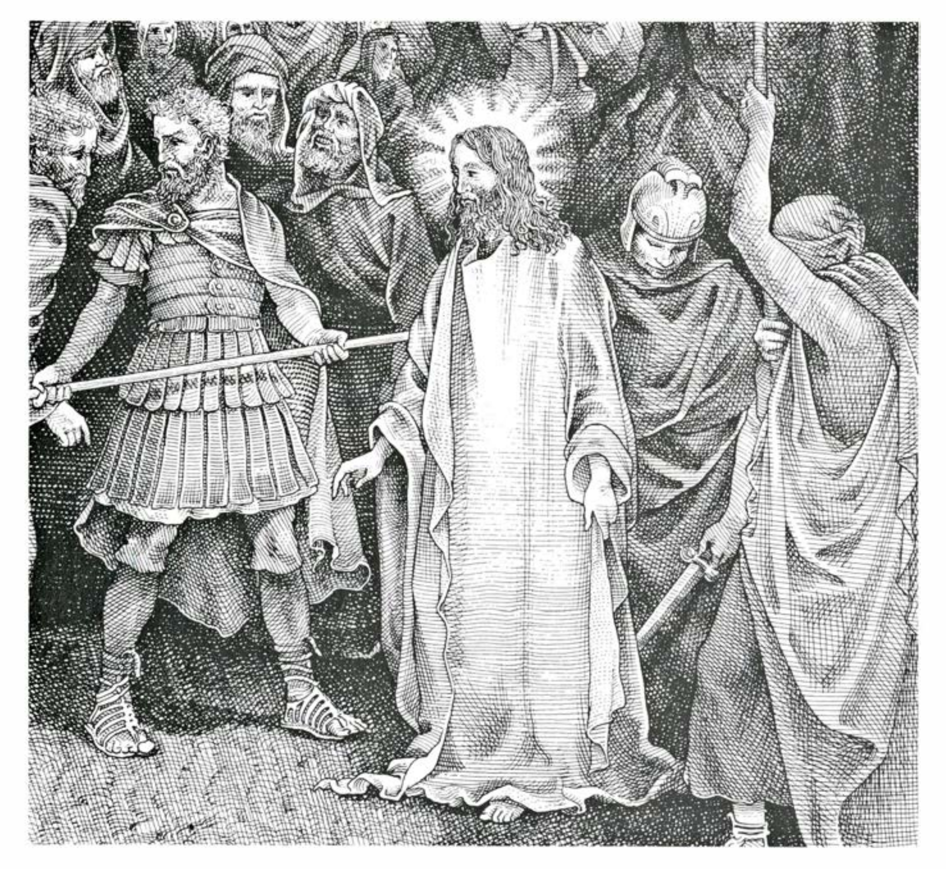






LA EVOLUCION DE UN SISTEMA ESTELAR HABITABLE depende del valor de la constante de interacción gravitatoria. En estos diagramas de Hertzsprung-Russell la banda diagonal es la secuencia principal, donde están secuenciadas las estrellas que se hallan en un estadio estable de su evolución. El diagrama superior muestra la evolución del Sol en el universo actual. El gas y el polvo de la nube protosolar se contrajeron bajo el influjo de su propia gravitación para formar el Sol, una estrella incluida en la secuencia principal. La contracción duró aproximadamente 10 millones de años. Tras permanecer 10.000 o 15.000 millones de años dentro de la secuencia principal, el Sol explosionará y se convertirá en una gigante roja. No mucho después evolucionará hasta hacerse una enana blanca y luego, lentamente, se irá apagando. El diagrama intermedio muestra cómo evolucionaría el Sol en un universo cuya constante de interacción gravitacional fuese de un orden de magnitud mayor. El Sol, en tal caso, pasa poco tiempo dentro de la secuencia principal y se convierte rápidamente en una gigante azul. La mitad de la derecha de la secuencia principal no tiene, o apenas tiene, estrellas. El diagrama inferior muestra cómo evoluciona el Sol en un universo cuya constante de interacción gravitacional es de un orden de magnitud más pequeño. El protosol entra en la secuencia principal como una enana roja, que permanece en la secuencia durante un tiempo larguísimo pero irradia poca energía. La mitad izquierda de la secuencia principal tiene escasas estrellas. Ni una gigante azul ni una enana roja podrían sustentar vida; la gigante azul se extingue demasiado deprisa y la enana roja irradia demasiado débilmente. El principio antrópico afirma que la presencia de vida en la Tierra explica por qué el Sol se halla en la línea divisoria entre las gigantes azules y las enanas rojas, y, en consecuencia, por qué la constante gravitacional tiene el valor que se le observa.





MUNDOS ALTERNATIVOS sugeridos por dos dibujos cuya base es un grabado de Gustave Doré. En un mundo Judas traiciona a Cristo con un beso; en otro mundo Cristo no es traicionado. Este ejemplo lo adujo Gottfried Wilhelm von Leibniz, hace unos tres siglos, en apoyo de su tesis de que hay un número infinito de mundos posibles, cada uno de ellos internamente consistente. Algunos de tales mundos tendrían leyes físicas, constantes fundamentales y condiciones iniciales muy diferentes de las del mundo de nuestra observación. Leibniz pensaba que este mundo que observamos se distingue de otros mundos posibles por tener la máxima variedad de fenómenos posible bajo las leyes físicas que describen los fenómenos. A Everett se debe la explicación mecánico-cuántica de la interpretación pluricósmica. El principio antrópico se asocia a una versión de la idea de los mundos alternativos. Para que un mundo sea observable ha de satisfacer una exigencia biológica: incluir rasgos apropiados para la aparición de la vida.

ción definida. Lo que ocurre durante una medición es que se selecciona un mundo de entre la infinita serie de posibilidades. La función de onda es todavía importante: sigue sirviendo para describir la totalidad de los mundos.

Por extravagante que parezca la interpretación pluricósmica, no puede ser descartada por pruebas físicas: es compatible con los resultados de los experimentos. Y tiene la virtud de conciliar la continuidad de la función de onda mecánico-cuántica con la discontinuidad de los procesos de medición.

El concepto de otros mundos no fue original de Everett. Unos tres siglos antes, Gottfried Wilhelm von Leibniz sostuvo que son infinitos en número los mundos posibles, cada uno de ellos intrínsecamente coherente y con sus características propias. Algunos de esos mundos diferirían grandemente del actual, tendrían unas condiciones iniciales muy extrañas para nosotros, así como las constantes fundamentales y las leyes de su naturaleza; otros mundos sólo diferirían en pocas cosas del que conocemos: por ejemplo, podría haber un mundo idéntico al nuestro salvo en que Julio César no cruzase el Rubicón. En otro mundo la diferencia consistiría en que Judas no traicionara a Cristo. La única restricción a un mundo posible es que no puede violar la ley de no-contradicción: no hay mundo alguno en el que César cruce y a la vez no cruce el Rubicón.

En la interpretación pluricósmica que hace Everett de la teoría cuántica, todos los mundos son igualmente reales. Para Leibniz, por otro lado, hay un principio de realidad que distingue a un mundo real de entre todos los demás mundos posibles. En opinión de Leibniz, la investigación científica revelaría que el mundo que observamos realiza al máximo una propiedad a la que él llamó en diversas ocasiones "economía", "perfección" y "optimidad". El último de estos términos es el más revelador. Sostenía Leibniz que el mundo óptimo exhibe la más rica variedad posible de fenómenos bajo las leyes físicas que describen los fenómenos. Empleó el concepto de optimidad para explicar las leyes de la reflexión y la refracción en óptica, y el mismo concepto le inspiró para desarrollar el principio de conservación de la energía.

Al combinar el principio antrópico con la interpretación pluricósmica de la mecánica cuántica, Carter introduce también un principio de realidad. La compleja propiedad por la que el mundo real se distingue no es la idea leibniziana de optimidad, sino una propiedad

a la que llamaré capacidad de sustentar la vida. Del infinito conjunto de mundos de Everett sólo considera Carter reales aquellos que satisfagan un requisito biológico: deberán incluir rasgos que hagan posible "la existencia de cualquier organismo descriptible como un observador".

Carter se basa en esta idea para explicar la debilidad de la gravitación. De acuerdo con la interpretación pluricósmica, tal vez existan mundos en los que la constante gravitacional tome todos los valores posibles, desde los muy débiles hasta los muy fuertes. El principio antrópico puede entonces explicar por qué vivimos en un mundo donde dicha constante tiene el valor observado. Carter demuestra que, si la constante gravitacional fuese muy diferente, los planetas o no se habrían formado o no habrían durado lo bastante como para que en ellos evolucionara vida inteligente. Siendo así que, presumiblemente, un observador requiere un planeta en el que habitar, la existencia de un observador se halla en estrecha dependencia del valor de la constante gravitacional.

La demostración de Carter se fundamenta en una interesante propiedad de las estrellas llamadas estrellas de secuencia principal, entre las que se incluye el Sol. Estas estrellas se mantienen en un estado estacionario de su evolución en el que la energía liberada por la fusión termonuclear contrarresta la fuerza de la atracción gravitatoria, equilibrándose ambas. Se las designa como estrellas de secuencia principal porque en un diagrama de Hertzsprung-Russell (que representa la relación entre la luminosidad y la temperatura superficial) forman una secuencia evolutiva dentro de una estrecha franja. La mayoría de las propiedades de las estrellas no dependen sensiblemente del valor de la constante de interacción gravitacional. Una excepción a esto es la nítida división de las estrellas de secuencia principal en gigantes azules (estrellas calientes, brillantes, de gran masa) y enanas rojas (estrellas frías, débiles, compactas). La luminosidad de una estrella es proporcional a la cuarta potencia de su masa, y así una gigante azul convierte rápidamente su sustancia en energía; su tiempo de vida es breve. Una enana roja libera, comparativamente, poca energía y tiene una existencia mucho más larga.

La emergencia de la vida requiere dos cosas de una estrella. Primera, que dure lo suficiente para que lleguen a formarse los organismos vivos. Segunda, que irradie energía bastante para calentar una región habitable del espacio, esto es, una región donde un planeta pueda tener una órbita estable. Ninguna gigante azul ni enana roja satisfacen ambas condiciones: las gigantes azules se consumen demasiado de prisa y las enanas rojas irradian demasiado débilmente. Lo que se necesita es una estrella del tipo de nuestro Sol, cuya posición en la secuencia principal es precisamente intermedia entre las gigantes azules y las enanas rojas; sólo una estrella así posee una combinación idónea de duración vital y poder radiante. Si la constante de interacción gravitacional fuese de un orden de magnitud mayor, la secuencia principal constaría toda ella de gigantes azules. Si fuese de un orden de magnitud menor, la secuencia principal constaría solamente de enanas rojas. En ninguno de los dos casos existirían estrellas capaces de originar y sustentar vida.

Como lo reconocía Carter, su argumento es más bien especulativo. La formación de los planetas no es aún lo bastante conocida como para que excluyamos por completo la posibilidad de que se formen planetas habitables en un universo cuya constante de interacción gravitacional sea diferente. Pero adviértase que esta incertidumbre afecta, no a la lógica del argumento, sino a sus premisas empíricas.

Carter se ha atenido al principio antrópico en otros contextos que están basados en premisas empíricas más sólidas. Por ejemplo, ha hecho notar que la constante de interacción asociada a la fuerza fuerte, o nuclear, "sólo periféricamente es lo bastante fuerte para contener [a los protones y neutrones] en los núcleos; si fuese un poco más débil, el hidrógeno sería el único elemento y probablemente sería también incompatible con la existencia de vida".

Collins y Hawking hubieron de llegar asimismo a invocar el principio antrópico. Su investigación se proponía explicar dos observaciones: la isotropía a gran escala del universo, y particularmente de la radiación de fondo de microondas, y la presencia, a escala menor, de inhomogeneidades, como las galaxias. Encontraron que los factores cruciales son la inicial velocidad de recesión de la materia creada en la gran explosión y la velocidad de escape de la materia (la velocidad que necesitaría ésta para superar su atracción gravitatoria). Si la velocidad de recesión fuese menor que la de escape, el universo se colapsaría antes de que pudiera desarrollarse la isotropía. Si la velocidad de recesión fuese mayor que la de escape, no podrían formarse las galaxias ni otras acumulaciones de materia, a no ser que, en el instante mismo de la gran explosión, hubiese inhomogeneidades de menor escala en la distribución inicial de la materia. Pero tales inhomogeneidades habrían dado por resultado anisotropía a gran escala en el universo presente. Collins y Hawking concluyeron, a su pesar, que la observada combinación de isotropía a gran escala y acumulaciones a menor escala sólo puede producirse si la velocidad de recesión es exactamente igual a la velocidad de escape. De lo que se infiere que el universo observado es ciertamente un universo privilegiadísimo, donde la velocidad de recesión tiene un valor decisivo entresacado de una serie infinita de posibilidades.

Collins y Hawking sugirieron que la chocante singularidad del universo observado podría entenderse gracias al principio antrópico. Empezaron por postular un conjunto numéricamente infinito de universos que tuvieran todas las condiciones iniciales posibles, incluidos todos los valores de la velocidad de recesión. En casi todos estos universos la materia no podría condensarse para formar galaxias. El único universo en el que la materia podría juntamente formar galaxias y exhibir isotropía a gran escala era un universo cuya velocidad de recesión fuese igual a la velocidad de escape. Collins y Hawking concluyeron que puesto que "la existencia de galaxias parece ser una condición previa necesaria para el desarrollo de cualquier forma de vida inteligente..., el hecho de que hayamos observado que el universo es isótropo, solamente es, por tanto, una consecuencia de nuestra existencia".

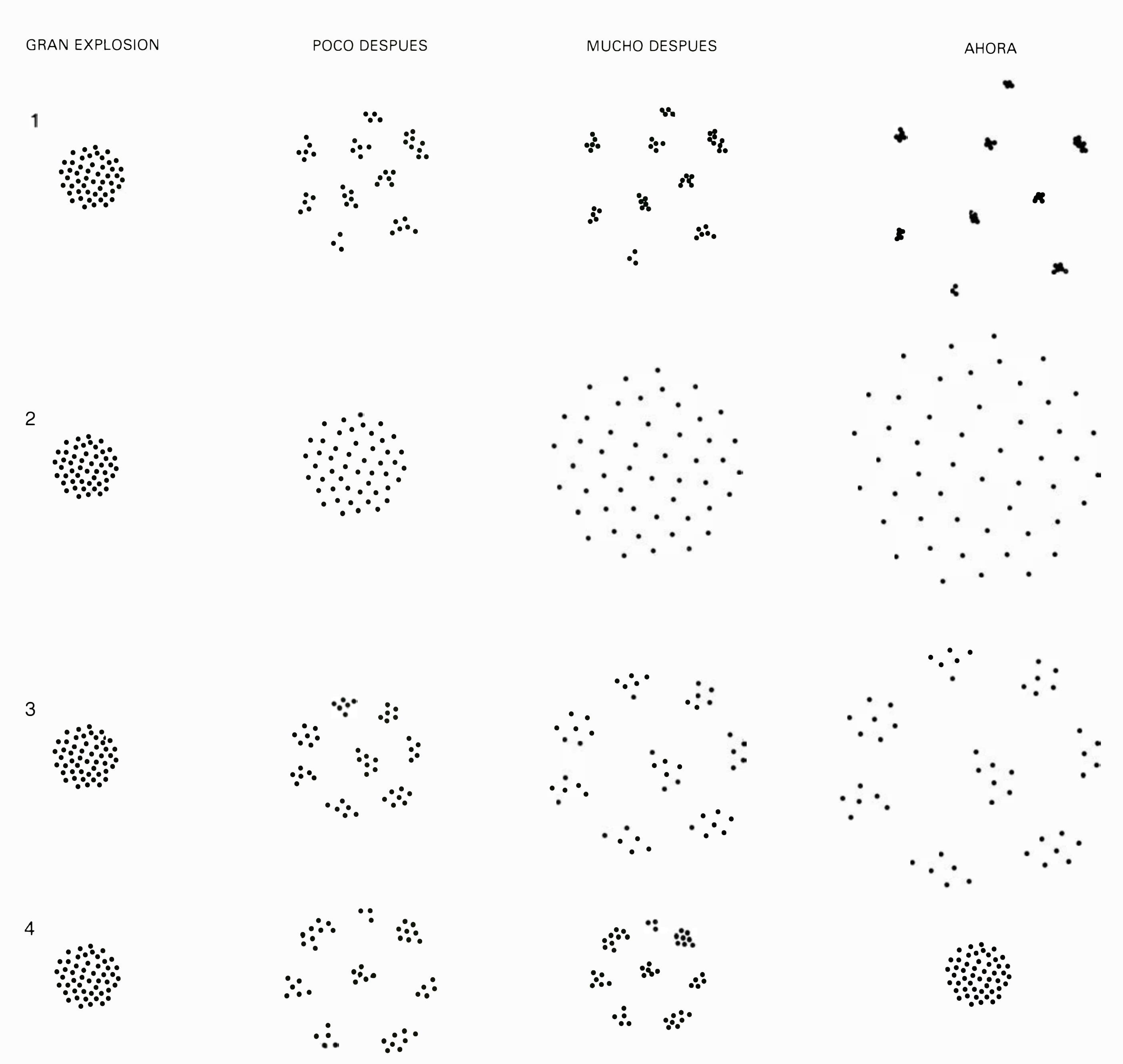
¿Qué sugiere el principio antrópico sobre la estructura total del mundo? Supóngase que en años venideros la línea de investigación antrópica revela que aun el más pequeño cambio de alguna de las condiciones iniciales del universo o del valor de cualquier cantidad fundamental no habría permitido la evolución de la vida. Esto sugeriría que, de todos los mundos posibles, el actual es el único apropiado para la vida. Sin embargo, harían falta más pruebas antes de que tal conclusión pudiera proponerse con toda confianza.

Wheeler ha hecho una pregunta aún más general: "¿Cómo se originó el universo?". La mayoría de los filósofos de la ciencia niegan que esta cuestión tenga científicamente sentido: toda respuesta requeriría, al parecer, un marco de referencia metacientífico, porque la estructura misma de la ciencia (a saber, el espacio-tiempo) y las leyes de la física que describen el espacio-tiempo se

originaron al crearse el universo en el instante de la gran explosión. A pesar de ello Wheeler sostiene que, en tanto no haya completa evidencia de la falta de sentido o indecidibilidad de la cuestión, no se puede uno contentar con "que asunto de tanta importancia quede siempre colgando en el aire, cual pelota de interminables jugueteos de indecisión".

Wheeler aborda la cuestión analizando la lógica de las explicaciones adoptadas en las teorías físicas desde la Revolución Científica del siglo XVIII. Mantiene que la lógica consiste en reducir cada fenómeno a otro más fundamental. Así, en la química, el concepto de valencia fue reducido al de propiedades eléctricas de los átomos, y la temperatura de un gas fue reducida al movimiento de sus átomos y moléculas. Parece ser que la lógica de la reducción sólo lleva a dos resultados posibles, de los que ninguno es, según Wheeler, aceptable: la teoría física podría terminar en algún indivisible objeto o campo fundamental; o, si no, la reducción podría ir descubriendo una capa estructural tras otra hasta el infinito.

Wheeler escapa de este dilema pro-



LOS MODELOS DE EVOLUCION COSMICA indican que el universo presente, caracterizado a la vez por la isotropía a gran escala y las inohomogeneidades a menor escala, tales como las galaxias, no habría surgido si la velocidad de recesión de la materia formada en la gran explosión no hubiese sido igual a la velocidad de escape de la materia (la velocidad que ésta necesita para contrarrestar su atracción gravitatoria). La primera línea de dibujos muestra la actual evolución del universo. La gran explosión tuvo arbitrarias perturbaciones a menor escala, las cuales se desenvolvieron en forma de inhomogeneidades que, condensándose, se convirtieron en galaxias. La segunda línea muestra la evolución de un universo en el que la velocidad de recesión es mayor que la velocidad de escape, y la gran explosión es homogénea. El universo presente sería completamente homogéneo y, por tanto, no habría en él galaxias. La situación no sería mejor en tal universo si la gran explosión hubiese tenido arbitrarias perturbaciones a pequeña escala, como lo indica la tercera línea. Esas perturbaciones darían por resulado inhomogeneidades a gran escala, que no se condensarían en forma de galaxias. La línea cuarta presenta la evolución de un universo donde la velocidad de recesión es menor que la velocidad de escape y la gran explosión tiene perturbaciones de menor escala. Las perturbaciones crecen hasta convertirse en inhomogeneidades que, andando el tiempo, empiezan a condensarse. Pero antes de que puedan formarse galaxias, el universo se colapsa. Parece ser que la observada combinación de isotropía a gran escala y agrupamientos a pequeña escala sólo se podría derivar de condiciones iniciales altamente específicas. C. B. Collins y Steven W. Hawking han propuesto que la singularidad del universo observado puede entenderse combinando el principio antrópico con la idea de mundos alternativos. Partiendo de un conjunto de infinitos universos que tengan todas las relaciones posibles entre la velocidad de recesión y la velocidad de escape, el único universo en el que podía emerger la vida es el universo actual.

poniendo que se acabe de una vez con ese modo de discurrir reductivo. "Desespérase uno y se pregunta si la estructura, más bien que terminar en algún objeto pequeñísimo o en algún campo fundamentalísimo, o que seguir de capa en capa, no llevará finalmente hasta el observador mismo en alguna suerte de cerrado circuito de interdependencias". Su argumento se inspira en la conexión establecida en la mecánica cuántica entre el observador y el fenómeno cuántico por él observado. La interpretación pluricósmica de la mecánica cuántica minimiza el papel del observador, porque en ella se considera que el mundo observado por éste no es más real que cualquier otro mundo. En cambio, las interpretaciones más comunes de la mecánica cuántica definen la realidad como aquello que es observado; el observador contribuye a la realidad por el acto mismo de observación. Wheeler adopta una versión extremada de esta idea al proponer que para que un universo sea real ha de evolucionar de tal modo que en él puedan llegar a existir observadores.

En apoyo de esta tesis cita Wheeler el principio antrópico. Afirma que "jamás se ha dado razón alguna de por qué ciertas constantes y condiciones iniciales tienen los valores que tienen, excepto la de que, de no ser así, algo como nuestra capacidad de observadores, según la conocemos, sería imposible". Se pregunta si no se podría "concebir, como Carter lo hace 'un conjunto de universos' tal que sólo en una porción muy pequeña del mismo fuesen posibles la vida y la conciencia. O inquirir si no sería en absoluto imposible que llegara a existir universo alguno como no se le garantizara el producir vida, conciencia y capacidad de observación en alguna parte y por algún breve tramo de tiempo en su historia futura". Wheeler rechaza el común parecer según el cual la vida y la capacidad de observación son meros accidentes en un universo independiente de los observadores, y sostiene en vez de esto que "la mecánica cuántica nos ha obligado a tomar en serio y a sopesar la opinión, directamente opuesta, de que tan esencial es el observador para la creación del universo como el universo para la creación del observador".

Con esta hipótesis ha llevado Wheeler el principio antrópico mucho más allá de los dominios de la lógica de la explicación: ha cruzado el umbral de la metafísica. Pocos científicos o filósofos de la ciencia se sentirían cómodos con su visión. Queda por ver si irán ganando aceptación las aplicaciones menos generales del principio antrópico.

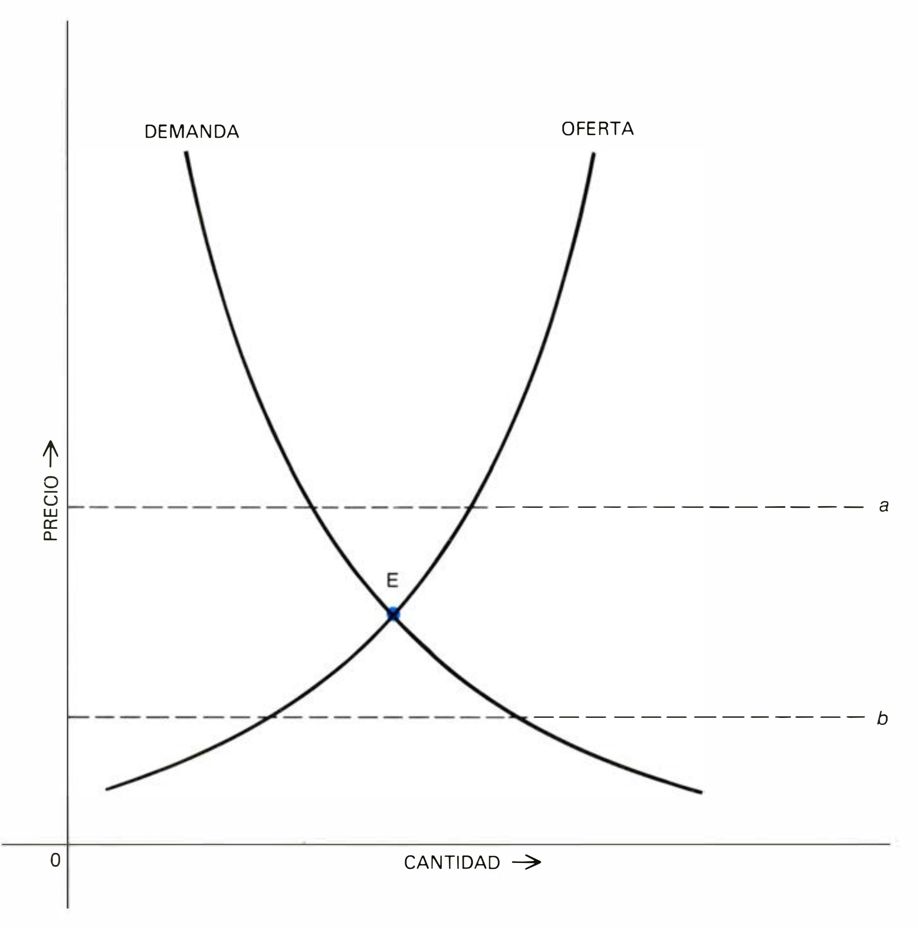
Juegos matemáticos

La curva de Laffer y otras gracias de nuestra economía de tipo mixto

Martin Gardner

No hay duda de que a los economistas les encanta trazar gráficas. Ya en los primeros decenios del capitalismo moderno, los economistas de la escuela clásica gustaban de explicar la teoría de precios construyendo gráficas de oferta y demanda como la que vemos abajo. Si el precio de una mercancía se encuentra al nivel indicado por la recta de trazos *a*, es fácil comprender, observando la intersección de la recta con las curvas, que la gente comprará menos dicha mercancía. Como el vendedor se encontrará entonces con exceso de oferta, tendrá que bajar el precio para poder colocar su producto. Si los precios se encuentran en el nivel inferior definido por la línea *b*, el aumento de demanda hará subir el precio y el vendedor habrá de producir más para atender la demanda.

Estas fuerzas de sentidos contrarios se equilibran en *E*, punto donde quedan compensadas mutuamente la demanda y la masa de producto puesta en venta. En este punto es donde el vendedor logra ganancia máxima. De mantenerse constante la oferta, un aumento general de la demanda provocará el desplazamiento hacia la derecha de la curva de demanda, y la altura del punto de equilibrio aumentará. Si se produce un aumento general de oferta, con demanda constante, la curva de oferta se desplaza hacia la derecha, y *E* cae. Si ambas curvas se trasladasen una misma cantidad hacia la derecha o la izquierda, *E* permanecería al mismo nivel.



Curvas clásicas de oferta y demanda

Estas curvas siguen siendo válidas, porque oferta y demanda desempeñan papeles de primera magnitud en todo sistema económico, incluso en aquellos donde no existe libertad de mercado; pero en nuestros días los economistas ya no se refieren a ellas con tanta frecuencia, porque en las economías de tipo mixto, como la nuestra, hay cientos de variables más que causan auténticos estragos en sus estilizadas formas. Valiéndose de ciertas estratagemas, los gobiernos consiguen que determinados productos tengan precios muy superiores o muy inferiores a los que alcanzarían en un mercado libre. Las organizaciones sindicales presionan a las empresas para conseguir de ellas mejores salarios, generando costos que luego las empresas cargan en los precios. Los oligopolios han descubierto sutiles procedimientos para resguardarse de las fluctuaciones del mercado, cosa que forzosamente deben hacer para seguir siendo eficientes.

En la década de los 60, estando en plena vigencia las teorías económicas keynesianas ("Hoy somos todos keynesianos", diría por entonces Richard Nixon), causó fuerte conmoción entre los economistas la llamada curva de Phillips. El primero en proponerla fue el economista londinense Alban William Housego Phillips, en 1958; su curva fue aplicada a la economía norteamericana en 1960 por dos economistas keynesianos, Paul A. Samuelson y Robert M. Solow. Como podemos observar en la ilustración de la página siguiente, las curvas de Phillips expresan gráficamente la relación de dependencia inversa entre desempleo e inflación. Tomando en cuenta la capacidad de las organizaciones sindicales y empresariales para influir en los precios, la curva de Phillips hace ver que los objetivos de alcanzar pleno empleo y al mismo tiempo precios estables son incompatibles en economías de tipo mixto. El pleno empleo (*F*) sólo puede lograrse a costa de continua inflación. La estabilidad de los precios (inflación nula) es imposible sin altas tasas de desempleo (*U*).

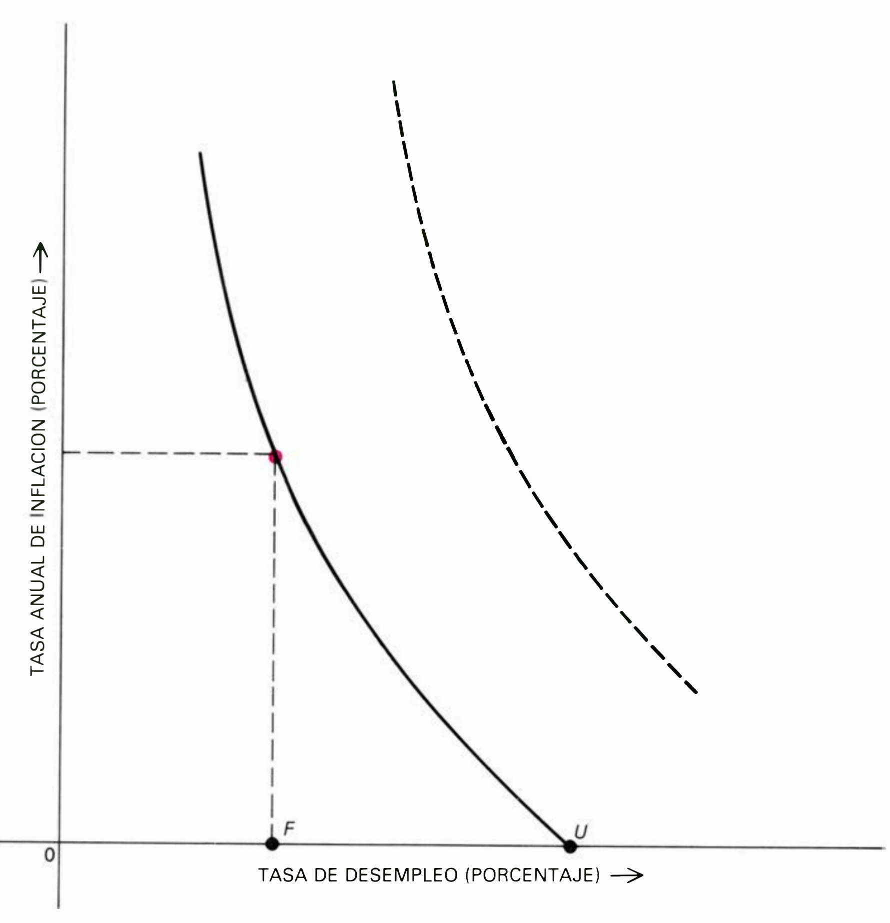
Así pues, ¿qué hacer? Lo mejor que podemos esperar, nos dicen las curvas, es ver de hallar una transacción razonable entre ambos casos extremos, que cause el menor daño posible. Si los precios se disparan, dejemos que un período de recesión los obligue a descender. Si tenemos demasiados parados, permitamos que la inflación genere nuevos empleos. Con un poco de suerte, los

gobiernos podrían atinar con un punto de la curva que combine una tasa "normal" de desempleo con una cuota de inflación "aceptablemente" baja, que pudiera estar en torno al 4 o 5 por ciento anual.

Y entonces ocurrió algo muy curioso. La economía, tanto en los Estados Unidos como en casi todo el mundo, cayó en el misterioso trance denominado "estanflación" ("estancamiento con inflación"), donde, contrariamente a lo previsto por la curva de Phillips, tanto inflación como desempleo crecen simultáneamente. Tal proceso hace deslizarse peligrosamente hacia arriba y hacia fuera el trazado de la curva, como indica la línea a trazos de aquella ilustración. La ensoñada "transacción", el razonable equilibrio entre paro e inflación, se ha transformado en una pesadilla, que los economistas han llamado "el cruel dilema". Los gobiernos carecen de "programas a la carta", de fórmulas económicas que no ahonden más la recesión o que no aceleren la actual inflación galopante.

Los keynesianos lucharon para rescatar la curva. Pronto se vio claramente que no existe nada parecido a una curva de Phillips que sea estable a corto plazo. La curva puede adoptar docenas de trazados, según cuales sean las variables (incluidas algunas de carácter psicológico, como las esperanzas y expectativas de los agentes económicos) que se tomen en cuenta, y además varía ampliamente de unos lugares a otros. ¿Existirá entonces alguna curva de Phillips que resulte estable a largo plazo? Hay opiniones para todos los gustos. E incluso, de existir, tampoco hay acuerdo en cómo tendría que aplicarse. ¿Deberían quizá los gobiernos tratar de que sus economías se deslizasen arriba y abajo de la curva, haciendo subir y bajar alternativamente paro e inflación, como dos niños en un balancín? ¿Deberían tal vez ensayar distintas y arriesgadas formas de "rodearla"?

Según los keynesianos, una fuerza, llamada tensión de la demanda, trata de estirar más la gráfica, llevándola hacia una vertical, mientras otra fuerza, la presión de los costos, intenta achatarla hacia la horizontal. Al cabo resultaría, según ellos, una curva intermedia fuertemente inclinada hacia abajo. Evidentemente, lo que se precisa es algún procedimiento para lograr que la curva se deslice hacia abajo y hacia la izquierda, permitiendo así transacciones que no aboquen al caos social. Ciertos economistas –John Kenneth Galbraith entre ellos– opinan que tan sólo podrá alcanzarse este objetivo combinando oportunamente las políticas monetaria y fiscal con un adecuado control de precios y salarios. Nada podría ser peor, dice Milton Friedman. Desde el enfoque monetarista de Friedman, la forma que a largo plazo tomaría la curva de Phillips sería una vertical situada en la tasa "natural" de desempleo; todo esfuerzo por alcanzar transacciones que pretendan reducir el desempleo a tasas inferiores a la natural desencadenará inflación galopante.



La curva de Phillips

La curva de Phillips, escribía Daniel Bell el año pasado (resumiendo anteriores comentarios de Solow), "ha generado más puestos de trabajo a los economistas... que cualquier proyecto de obras públicas desde la construcción del Canal del Erie". Hoy, la curva de Phillips se está convirtiendo rápidamente en desvaído símbolo de una idea, a saber, que paro e inflación no son calamidades independientes, sino que están ligadas por complejas relaciones funcionales que nadie es hoy capaz de comprender.

Actualmente, y como consecuencia del creciente interés que suscitan las teorías de oferta, la curva de moda es otra enteramente distinta; se trata de la curva de Laffer. A sus 41 años, Arthur B. Laffer es profesor de ciencias empresariales en la Universidad de California del Sur. Su curva fue bautizada y dada a conocer por Jude Wanniski, redactor que fue de *The Wall Street Journal*, en un libro que se ha convertido en la biblia de todos los "ofertistas", titulado nada menos que *The Way the World Works: How Economies Fail and Succeed* ("Así funciona el mundo: Sistemas económicos que fracasan... o triunfan" sería la traducción del título), publicado por Basic Books en 1978. La ilustración de la página siguiente nos muestra cómo enfoca Wanniski el significado de la curva de Laffer, al comienzo del capítulo sexto de su libro.

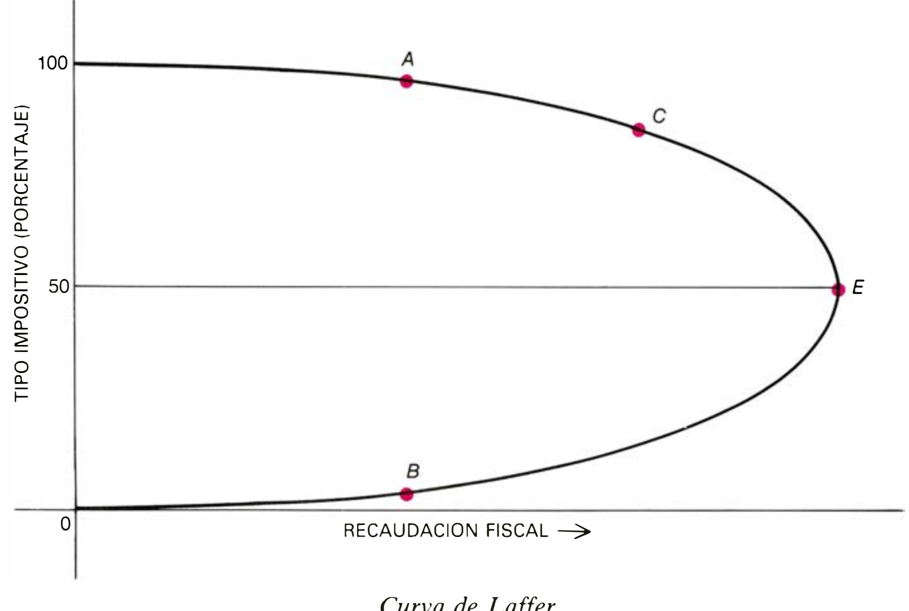
¿Acaso no es la imagen misma de la belleza? Como cualquier niño puede comprender sin más que fijarse en el extremo inferior de la curva, si el gobierno reduce a cero sus tasas de imposición, nada recaudará. Y si eleva sus tipos fiscales hasta el 100 por ciento de los ingresos de cada contribuyente, tampoco recaudará nada. ¿Por qué? Porque en tal caso nadie se molestaría en trabajar para recibir renta. En efecto, si todos sus ingresos se los llevase el estado, los individuos retornarían a una economía de trueque: los pintores sólo le pintarían la casa a los dentistas si éstos les arreglasen la boca.

La curva de Laffer se hace más interesante conforme nos deslizamos por sus brazos, hacia el centro. En el punto *A*, donde los impuestos no alcanzan a ser del 100 por ciento, algunas personas empiezan a encontrar conveniente recibir parte de sus ingresos en forma de renta imponible. En el punto *B*, las ruedas de la economía sisean alegremente, la producción, sin freno apenas, se mantiene a niveles elevados, pero a causa de lo reducido de los tipos impositivos, el gobierno recauda igual de poco que con los duros impuestos de *A*.

Fijémonos ahora en el punto *E*, situado en el vértice derecho de la curva. En este punto se encuentra la tasa fiscal que hace máximos los ingresos del gobierno. Si la tasa cayera por bajo de *E* tal vez la producción recibiera algún estímulo, pero el gobierno se vería debilitado. Como por definición *E* es el punto de recaudación máxima, el gobierno recibe también menores ingresos si eleva los impuestos por encima de *E*. Los partidarios de la economía de oferta hacen notar que hay muchas razones para que así suceda. Para algunas personas ricas resulta entonces contraproducente seguir trabajando tan productivamente como antes. Otros huyen de las excesivas cargas fiscales y congelan su dinero en "refugios" improductivos. Algunos prefieren incluso emigrar a países cuya fiscalidad no sea tan onerosa. Si el gobierno precisa imponer fuertes cargas fiscales para programas de seguridad social, habrá millones de personas que no sientan incentivo alguno para trabajar lo más mínimo. ¿Para qué molestarse en hacerlo, si podemos lograr ingresos comparables del protector estado? Las grandes sociedades anónimas dedican partidas menores a los programas de investigación y desarrollo de nuevos productos. Los empresarios, columna vertebral del crecimiento dinámico, estarán menos dispuestos a correr riesgos. A consecuencia de estos y otros factores, las fuerzas económicas se hacen indolentes, y los ingresos fiscales decrecen.

Es importante comprender, nos dice Wanniski, que *E* no se encuentra forzosamente en el nivel del 50 por ciento, aunque bien pudiera estarlo. La forma de la curva de Laffer cambia, evidentemente, con las circunstancias. Así, en tiempos de guerra, en momentos en que tanto la población como el empresario están persuadidos de la necesidad de sacrificios, es posible imponer tasas fiscales muy elevadas y mantener al mismo tiempo la producción a pleno rendimiento. En la paz, por el contrario, todos somos menos altruistas.

Ahora bien, la esencia del razonamiento de los "ofertistas" es la convicción de que actualmente la economía norteamericana se encuentra en las proximidades de *C*, demasiado arriba sobre la curva de Laffer. Al rebajar los impuestos (lo que según algunos de estos economistas conlleva fuertes recortes en gastos de seguridad social) el brazo productor de la economía recibirá una inyección fuertemente estimulante, tanto que, en los Estados Unidos, pronto la economía irá deslizándose cuesta abajo por la curva de Laffer hasta el punto *E*, tal vez no inmediatamente, pero sí pronto. Los ingresos fiscales llegarán finalmente a poder financiar los crecientes gastos militares, la estanflación terminará, recomenzará el crecimiento dinámico, el presupuesto quedará equilibrado hacia 1984 y el sueño americano recobrará su lustre.



Curva de Laffer

Evidentemente, oferta y demanda están siempre entrelazadas; empero, los ofertistas se llaman a sí mismos de esta forma para subrayar sus diferencias con los neo-keynesianos. John Maynard Keynes resaltó la importancia de mantener niveles de demanda suficientes; para ello se dictan leyes de salario mínimo y se crean subsidios con cargo a la seguridad social. Los lafferistas han vuelto el razonamiento del revés, cargando el acento en la importancia de estimular la oferta. De no ir el gobierno a hombros del empresariado, la producción remontaría el vuelo, se desarrollarían nuevos inventos, se necesitarían nuevos puestos de trabajo y los salarios reales irían en alza. Toda la sociedad, y especialmente sus capas más pobres, saldría favorecida conforme la prosperidad creada en lo alto fuera descendiendo, empapándola.

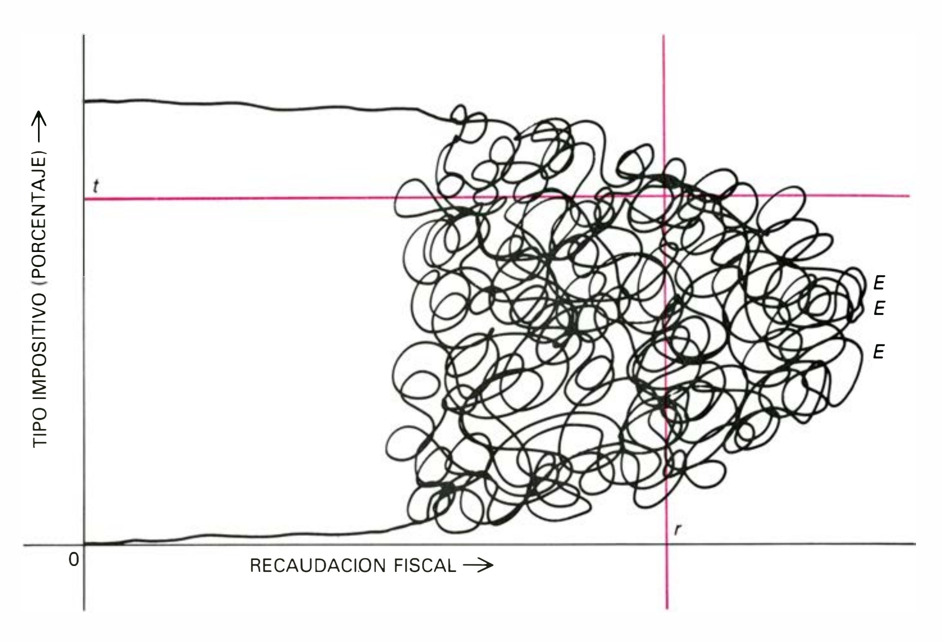
El segundo de los libros en proclamar las virtudes del lafferismo es *Wealth and Poverty*, de George Gilder, (Basic Boods, 1981). Su título parafrasea intencionadamente el famoso *Progress and Poverty* de Henry George, que causó conmoción a fines del siglo pasado al recomendar la abolición de todo impuesto, salvo uno, que gravaría la posesión de tierras. El libro de Gilder es más radical que el de Wanniski. "¡La regresión del impuesto ayuda a los pobres!", escribe Gilder en su libro. En cierta ocasión, William Safire describió el capitalismo como "el bien que puede resultar de la codicia". Gilder se enfurece cuando se dicen tales cosas; según él, el capitalismo está motivado por el bien que se origina al "dar". Pretende de esta forma decir que la mejor manera de dar a los pobres lo que quieren y necesitan, en especial a los jóvenes parados de los grupos minoritarios, es permitir que el mercado funcione libremente, sin cortapisas, para lograr así que el sistema económico recobre su dinamismo.

Lo malo de la curva de Laffer es que, al igual que la curva de Phillips, resulta demasiado sencilla para poder prestar servicio alguno, aparte su valor expresivo como símbolo de una idea. En el caso de la curva de Laffer, la idea de que la fiscalidad excesiva es contraproducente, es, además de una verdad palmaria, bastante antigua. El verdadero problema consiste en definir *cuánto* es *excesiva*. Ningún economista tiene la menor idea sobre el verdadero aspecto de la curva de Laffer, salvo en la vecin-

dad de sus extremos. Y aun cuando los economistas lo supieran, ignorarían en qué punto deberían situar la economía. Así, Irving Kristol, al defender un modelo económico "de oferta" en *Commentary* (abril de 1981) escribe que, aunque no sabe decir con exactitud en qué punto de la curva de Laffer nos encontramos, sí está seguro de que es "demasiado arriba". Y la disminución de impuestos que propone el presidente Reagan es precisamente lo que conviene para lograr que la economía se deslice hacia el punto *E*.

Con el fin de mejor acomodar la curva de Laffer a las complejidades de una economía de tipo mixto, como la nuestra, dominada por lo que Galbraith gusta en llamar "la tecnoestructura", y de tener en cuenta otras variables que la deforman, he ideado la que llamo "curva neo-lafferiana", por brevedad, "curva NL". (Véase mi artículo "Changing Perspectives on the Laffer Curve", en *The British Journal of Econometriciousness*, vol. 34, n.º 8, págs. 7316-7349; agosto de 1980.) Vemos la curva NL en la ilustración de esta página. Notemos que, en la vecindad de sus extremos, esta preciosa curva recuerda de cerca la curva de Laffer "clásica", demostrando así que la vieja curva no era tan sólo una primera aproximación carente por completo de valor. Sin embargo, conforme la curva se interna en las complejidades del mundo real, se ve arrastrada y envuelta por lo que yo llamo "tecnomaraña". En esta región, el trazado de mi curva se funda en un complejo análisis estadístico (debido a Persi Diaconis, del departamento de estadística de la Universidad de Stanford) realizado a partir de los más fiables datos económicos relativos a la economía norteamericana, a lo largo de 50 años. En el gráfico real, los datos quedan representados por un enjambre de puntos densamente apiñados, por lo que la forma real de la curva es hasta cierto punto arbitraria. No obstante, mi curva sí permite ver con claridad meridiana cierto número de ideas importantes.

Tomemos un valor *r* cualquiera del segmento del eje de "recaudación fiscal" situado directamente debajo de la "tecnomaraña". La recta vertical trazada por *r* corta a la madeja en multitud de puntos; las ordenadas de estos puntos representan valores del eje de tasas impositivas con máxima probabilidad de producir la recaudación *r*. Debe observarse que también sucede así para el valor máximo de *r*, que determina a su vez varios puntos *E* sobre la tecnomaraña. Con dos palabras: hay más de una



*La curva neo-lafferiana (***NL***)*

tasa impositiva capaz de producir la recaudación máxima posible.

Fijémonos también en un valor cualquiera *t* interior al segmento del eje de tasas situado directamente a la izquierda de la madeja. La recta horizontal trazada por él corta también en multitud de puntos a la maraña; estos puntos corresponden a los valores del eje de recaudación que se obtendrán con mayor probabilidad al fijar la tasa de fiscalidad en *t*.

Observemos ahora que, en ciertos puntos de intersección, al reducir la presión impositiva se consiguen menores ingresos, mientras en otros ello conlleva aumentos de recaudación. Aun admitiendo que supiéramos en qué punto debe situarse nuestra economía, no está claro, en vista de la maraña, cuáles deberían ser las políticas monetaria y fiscal que más rápidamente conducirían nuestra economía hacia el punto *E* más cercano.

Al igual que la antigua curva de Laffer, el valor de la nueva es sólo metafórico, aunque sin duda ofrece un modelo más perfecto de la situación real. Puesto que es reflejo estadístico de conductas humanas, su forma cambia constantemente (lo mismo que la curva de Phillips) y de manera impredecible. La curva quedaría mejor representada por una película cinematográfica que captara su versátil y mudable carácter. Mas como la recogida de datos requiere tanto tiempo, y el análisis estadístico de sus parámetros más todavía, para cuando se logra dibujar una de estas curvas NL ya está caduca y, por tanto, no es de ninguna utilidad. No obstante, por intermedio de uno de los más fidedignos informantes de Jack Anderson he podido enterarme de que la Fundación Smith Richardson ha financiado secretamente un proyecto de coste multimillonario (en dólares) en el Stanford Research International, cuya finalidad es la construcción de curvas NL. Tal vez valiéndose de programas más eficientes y con la ayuda del velocísimo ordenador CRAY-1 del Lawrence Livermore Laboratory sea posible asignar valores de probabilidad puestos al día a los distintos puntos de intersección. Si semejante objetivo pudiera conseguirse, la curva NL podría llegar a ser valioso instrumento de predicción que permitiera a la Administración tomar decisiones racionales en materia económica.

Los lafferistas están dando prueba de la más suprema confianza en sí mismos, adornada con un superlativo desprecio por las opiniones de sus detractores. De los 18 economistas que han sido distinguidos con el premio Nobel, tan sólo dos, Milton Friedman y Gunnar Myrdal, figuran en el índice de *The Way the World Works*. Ni siquiera Alan Greenspan, hoy considerado de la "vieja derecha", ya abandonada, les merece una mención. Dado que tanto Friedman como Wanniski son mentores del actual conservadurismo económico, podría suponerse que Wanniski habría de tener elevada opinión de las teorías de Friedman. No es así. En su libro, Wanniski explica prolijamente por qué tres famosos modelos económicos –el marxista, el keynesiano y el friedmaniano– son todos ellos erróneos. Ni siquiera consiguen explicar la catástrofe económica de 1929.

Se dispone hoy de una ingente biblio-

Solución al problema de Scott Kim, con 21 reinas

grafía relativa a las muchas causas de aquella crisis, en gran parte escrita por economistas eminentes. Podemos tirarla a la basura. Wanniski ha dado con la verdadera razón. Nada malo le hubiera pasado al mercado de valores si Herbert Hoover le hubiera dejado seguir su propio curso. En lugar de ello, Hoover y el Congreso cometieron un craso error. Escribe Wanniski: "El hundimiento de la Bolsa, en 1929, y la subsiguiente Gran Depresión sobrevinieron a causa de la aprobación (por las Cámaras) del Acta de Aranceles Smoot-Hawley, en 1930".

¿Y cómo es que la crisis financiera se produjo en octubre del año anterior? Muy simple. El mercado bursátil vio venir las horrendas consecuencias de las restricciones al libre comercio que estaban en gestación. No todos los ofertistas están de acuerdo en este punto. Uno de quienes sí lo están es Jack Kemp, coautor del proyecto de ley Kemp-Roth sobre fiscalidad (que preparó el camino al programa fiscal de Reagan). En su estimulante libro *An American Renaissance: A Strategy for the 1980s* (Harper & Row, Inc., 1979) Kemp nos asegura que Wanniski ha "demostrado más allá de toda duda razonable" la veracidad de su notable descubrimiento.

Mas, ¿qué opinan de las teorías de la oferta los economistas profesionales? A muchos de ellos, incluidos los más conservadores, les merece aproximadamente igual juicio que las teorías de Immanuel Velikovsky merecen a los astrónomos. Para Galbraith se trata de "una forma de timo razonablemente refinada". Walter W. Heller las ha comparado con un analgésico y Solow las califica de "purgante". El vicepresidente Bush las ha asimilado al vudú. Se dice, empero, que los libros de Wanniski, Gilder y Kemp pesan fuertemente en la actual Administración.

Los lafferistas parecen disfrutar prodigándose elogios mutuos. Laffer, héroe del libro de Wanniski, es citado en la contracubierta de la edición en rústica, y en su boca se ponen estas palabras: "Con toda sinceridad, opino que éste es el mejor libro de economía jamás escrito". En la cubierta, Kristol es más comedido. En su opinión, se trata "del mejor compendio de economía desde Adam Smith". Gilder asegura que Wanniski "ha alcanzado de la noche a la mañana una influencia y un prestigio de proporciones keynesianas". Gilder, por su parte, ha sido aclamado con euforia similar. David Stockman, director de presupuestos del presidente Reagan, ha ensalzado *Wealth and Poverty*, calificándolo de "prometeico por su vigor intelectual, por su penetración y perspicacia. Consigue hacer tambalearse de una vez por todas las entelequias del keynesianismo, las ilusiones del estado protector que agobian la fallida sabiduría económica de nuestra era."

¡Cuán perplejo debe encontrarse el presidente ante el violento encontronazo entre Friedman, viejo amigo suyo, y sus actuales consejeros lafferistas! (El encontronazo no sólo se debe a las opiniones de Friedman sobre política monetaria, sino también por el desagrado que produce a Friedman la "locura del oro", que ha hecho presa entre los ofertistas, quienes reclaman el retorno inmediado al "patrón oro".) En la sección de economía del *New York Times* (domingo, 26 de julio), Wanniski lanza sobre Friedman un ataque verdaderamente corrosivo. Su principal argumento es que si bien Friedman "apenas alcanza a medir metro y medio", su "peso" es tal que hoy es un "enorme lastre" que gravita sobre las espaldas de Menachem Begin, Margaret Thatcher, Ronald Reagan y la economía norteamericana.

¿Tendrán éxito las ideas lafferistas hoy en boga en la Administración norteamericana, o por el contrario, como muchos economistas temen, sumirán más profundamente al país en la inflación galopante y el desempleo? La verdad es que los economistas lo ignoran. La tecnomaraña está demasiado embrollada. Tal vez las clases ricas y ociosas decidan no invertir lo que van a ahorrarse de impuestos, como predicen los lafferistas que harán, sino gastarlo en bienes de consumo. Tal vez las clases trabajadoras opten por no rendir más, sino menos. Y quizá las grandes sociedades y grupos de empresas no hagan gran cosa con los excedentes fiscales, aparte de adquirir nuevas compañías.

Evidentemente, los ideólogos de todos los credos están convencidos de saber exactamente cómo responderá la economía al combinado de monetarismo y lafferismo que ha puesto en práctica la Administración. En efecto, tan vasta es la confianza que tienen en sí mismos, tan habilidosa es su capacidad para racionalizar y justificar sus fracasos, que a corto y medio plazo es imposible imaginar una escena económica capaz de probar la falacia de sus respectivos dogmas. El fracaso de una cualquiera de sus predicciones podrá siempre atribuirse a improvisadas decisiones del poder político o a imprevisibles acontecimientos internacionales. Así, es inconcebible que Friedman pueda considerar que la tasa inflacionaria de Israel –que ya se escribe con tres cifras–, los recientes disturbios de Gran Bretaña o la posibilidad de altísima estanflación en los Estados Unidos en 1983 puedan conllevar el menor desdoro para sus convicciones monetaristas, no obstante haber respaldado entusiásticamente Friedman las políticas de Begin, Thatcher y Reagan, y no obstante haber sido estos tres gobernantes fuertemente influidos por la forma de monetarismo que propugna Friedman.

Los lafferistas, por su parte, tienen todo tipo de vías de salida en caso de que la política de Reagan lleve al desastre. Algunos culparán a Friedman. A otros les bastará el plan de fuga que ya tiene trazado William F. Buckley. Aunque se ha dicho que las reducciones fiscales y presupuestarias de esta Administración son las mayores de la historia de los Estados Unidos, Buckley considera que ninguna de ellas es suficiente. "Lo malo de los recortes fiscales propuestos por Reagan, escribía Buckley en *National Review* (24 de julio), es que (*a*) son insuficientes, e insuficientemente programados; y (*b*) las reducciones en el presupuesto son igualmente insuficientes... No puede pretenderse alcanzar reducciones de importancia limitando los esfuerzos a una de las tres secciones del presupuesto. Es como decir que se quiere adelgazar y hacer ejercicio sólo con la pierna derecha."

Es de esperar que el presidente Reagan no trate ahora de reconciliar los enfrentados puntos de vista de estos conservadores recurriendo a la astrología. Semejante idea no es tan descabellada como podría pensarse. En una entrevista con Angela Fox Dunn, el presidente dijo seguir el consejo diario correspondiente a su signo zodiacal del horósco-

po que, distribuido por agencias de prensa, prepara Carroll Righter. Nacido un 6 de febrero, Reagan pertenece al signo de acuario. "Estoy convencido, le dice a Dunn, que se descubrirá que el 80 por ciento de los personajes del Hall of Fame de Nueva York son acuario."

Durante muchos años, el presidente Reagan y su esposa Nancy han sido amigos personales tanto de Righter (quien también aconseja a Gloria Swanson y a otros personajes de Hollywood) como de la astróloga Jeane Dixon, hoy en Washington. "No se me considera consejera suya, le dice sibilinamente Dixon a Warren Hinckle (*San Francisco Chronicle*, 19 de julio de 1980), pero alguno le doy." Joyce Jillson, que escribe una columna de astrología para agencias de información, publicada en el *Chicago Tribune*, y que cuenta entre sus clientes a varios estudios de Hollywood y diversas corporaciones multinacionales, declaró que, el año pasado, miembros del equipo de Reagan le pagaron 1200 dólares por los horóscopos de ocho hipotéticos candidatos a vicepresidente. No obstante, el jefe de la oficina de prensa de la Casa Blanca la ha llamado mentirosa. Michael Kramer escribe (*New York*, 21 de julio de 1980): "Ronald Reagan, dice Ronald Reagan, es un hombre agradable y bien intencionado, que adora a su familia, a quien le gusta consultar su horóscopo antes de tomar decisiones importantes, y que llora cuando ve *La casa de la pradera*."

¿Buscará ahora el presidente la ayuda del zodíaco para decidir si hacer caso de Friedman, de Laffer o de algún otro? No podemos estar seguros. Como ha escrito el economista William Nordhaus, de la Universidad de Yale (*The New York Times*, 9 de agosto): "Tan sólo podemos confiar en que la economía de la oferta se parezca más a la aspirina que a la talidomida".

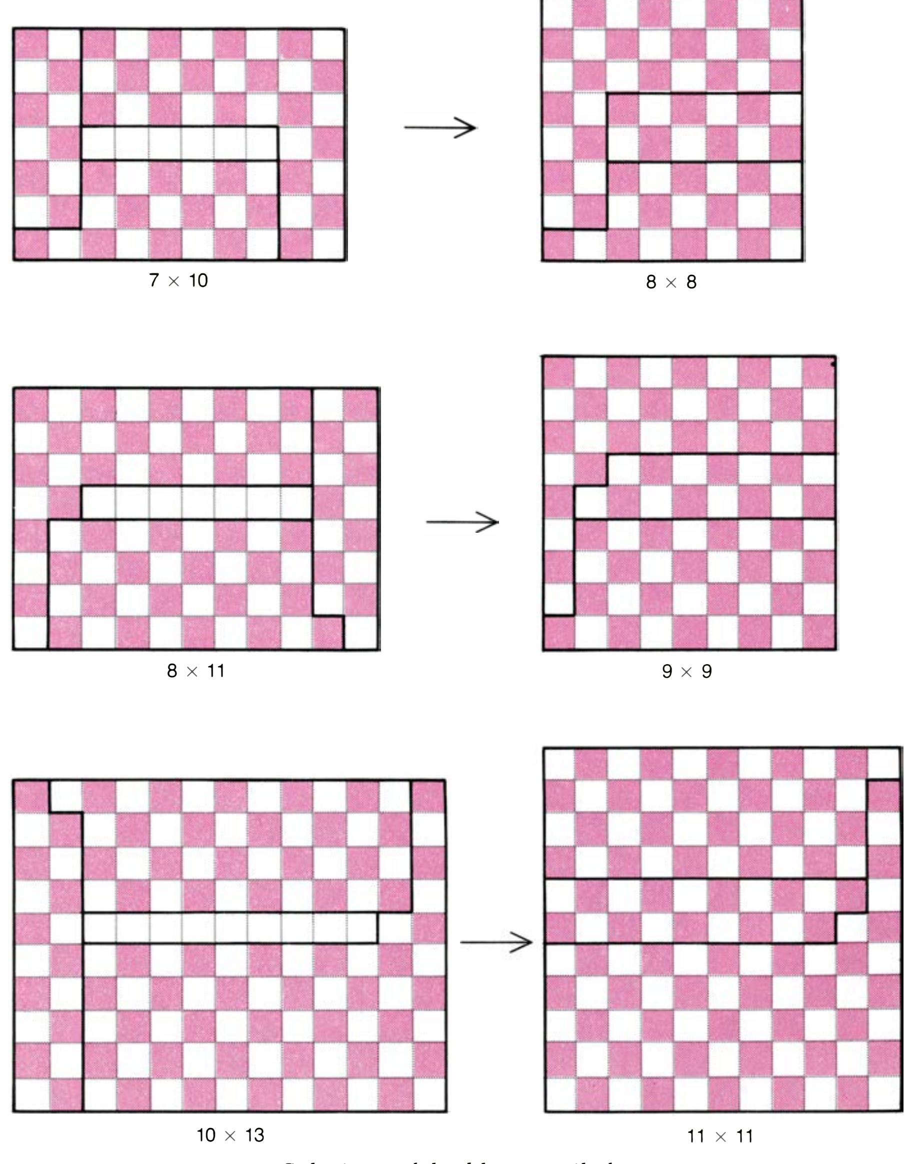
En agosto pasado mostré ya cómo doblar oblicuamente una hoja de papel según cuatro lados y producir una forma carente de planos de simetría, que sea, sin embargo, superponible a su imagen especular. Paul Schwink, de Carlisle, Iowa, y Piet Hein, de Copenhage, me señalaron que la solución expuesta es innecesariamente complicada. Puede conseguirse el mismo resultado plegando únicamente un par de lados opuestos, uno hacia arriba y otro hacia abajo.

En agosto di la que entonces era la mejor solución conocida del problema de Scott Kim, consistente en situar el máximo número de reinas sobre un tablero de ajedrez, de forma que cada una amenace exactamente a otras cuatro. Jeffrey E. Spencer fue un paso más allá que Kim, y descubrió una solución con 21 reinas. Esta inesperada disposición puede verse en la página anterior. ¿Será maximal? Nadie lo sabe.

Escribiendo sobre la parábola, el pasado mes de octubre, dije que al bajar el chorro de una manguera, el punto de altura máxima iba describiendo una parábola. No es así. La curva que sí es parabólica es la envolvente de los chorros.

Muchos lectores han confirmado una conjetura presentada en el número de octubre, a saber, que, con las actuales reglas gregorianas para calcular la fecha de la Pascua de Resurrección, la fecha más frecuente es el 19 de abril. Algunos de ellos llegaron a tal conclusión razonando a partir de los datos disponibles; otros utilizaron programas de ordenador para ensayarla en períodos de tiempo extremadamente grandes. Todos encontraron que el 19 de abril es la fecha más frecuente, y el 22 de marzo, la más excepcional.

En diciembre se pidió a los lectores que cortasen un rectángulo ajedrezado con un agujero oblongo en su interior, en tres piezas que puedan recombinarse (se permiten transformaciones de simetría) y formar un tablero de ajedrez ordinario coherente con el coloreado. En la ilustración superior, de esta misma página, vemos cómo conseguirlo. La generalización a cuadrados de número par de cuadros por lado es evidente. Para cuadrados de lado impar igual a n es preciso distinguir los casos $n = 1 \pmod 4$ y $n = 3 \pmod 4$. Vemos ejemplos de cada caso en las ilustraciones central e inferior. De nuevo es fácil comprender cómo se generalizan estas soluciones.



Soluciones del tablero mutilado

Taller y laboratorio

Meditaciones en torno a las burbujas formadas en una botella de cerveza

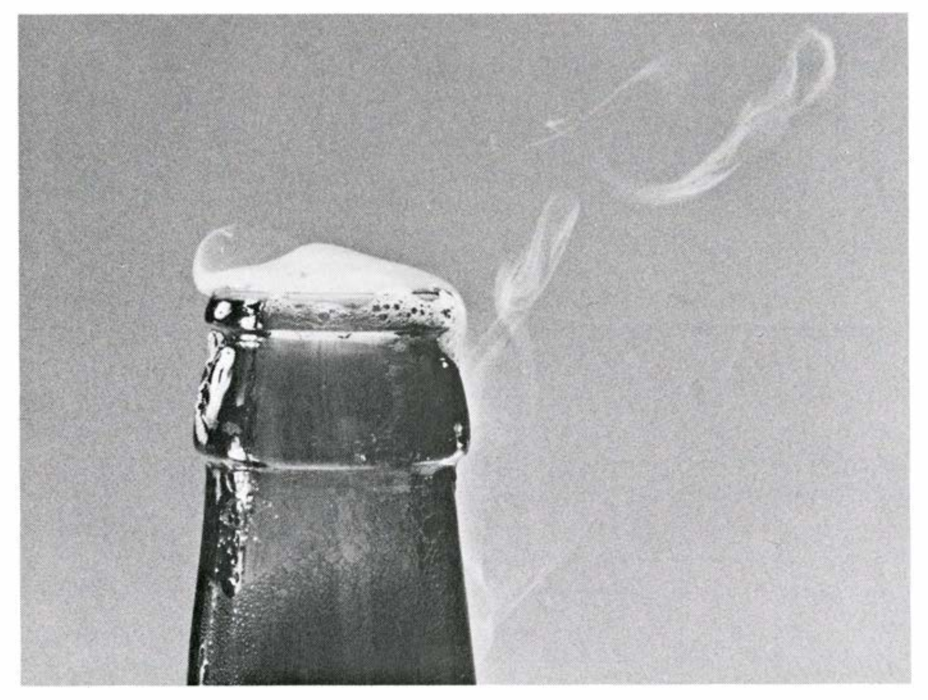
Jearl Walker

La efervescencia de la cerveza y demás refrescos carbónicos se debe al anhídrido carbónico que queda disuelto en el líquido, una vez embotellado, o enlatado, a presión. Sin ese anhídrido carbónico disuelto, el gusto y olor de la bebida serían muy diferentes. Cuando se abre el envase, la presión interior se libera y el anhídrido carbónico sale lentamente de la solución en forma de gas.

Al destapar una botella de cerveza, la variación brusca de presión que sufre el gas que está en contacto inmediato con la cerveza provoca una turbulencia suave y cavitación (o formación de vacíos parciales) en la superficie; ello da lugar a la aparición de burbujas de anhídrido carbónico, que revientan enseguida. El cambio de presión genera, asimismo, un leve vaho en la zona contigua a la boca de la botella abierta. Este vaho lo forman dos componentes: grandes gotas, que la cavitación lanza hacia arriba, y otras gotitas, menores, que se condensan y separan del gas saliente (aire, vapor de agua y anhídrido carbónico).

Para mí, la condensación se manifiesta mejor cuando la botella de cerveza está ligeramente fresca. La presión que reina en el vapor de la zona inmediata al líquido es entonces mayor que si la botella estuviera fría. Tan pronto se abra la botella, esa bolsa de gas se dilatará bruscamente sobre la atmósfera inmediata al cuello de la botella. Tal dilatación de gas consume energía, dado que la bolsa de gas ha de abrirse camino hacia la atmósfera: el gas debe realizar un trabajo aplicando una presión que desplace una superficie de separación (entre el gas y la atmósfera).

Como la expansión es muy rápida, no da tiempo a que se transfiera al gas energía de la cerveza, ni de la botella o la atmósfera: el gas va impulsado por la energía cinética de sus propias moléculas. Estas se encuentran animadas de un movimiento caótico, al que corresponde una energía cinética cuya medida indirecta es la temperatura del gas.



Vaho que se genera en una botella de cerveza recién abierta

Cuando las moléculas pierden energía, se mueven más despacio y baja la temperatura del gas. Esta disminución de temperatura hace que el vapor de agua se condense en gotitas, formando el vaho que se acumula momentáneamente sobre la boca de una botella de cerveza recién abierta.

Cuando se vierte en el vaso, la turbulencia y cavitación son mucho mayores y el líquido atrapa aire, que forma burbujas. En las zonas donde la turbulencia hace que disminuya la presión de la cerveza, las moléculas de anhídrido carbónico disueltas se reúnen para formar burbujas, que crecen a medida que van recogiendo más anhídrido carbónico. Por ser la densidad de las burbujas mucho menor que la de la cerveza, se elevan hasta la superficie, originando allí la consabida masa espumosa.

Una broma pesada muy corriente consiste en agitar, subrepticiamente, un envase cerrado de cerveza, o de cualquier otro refresco, y ofrecérselo luego a alguien. Cuando el embaucado de turno lo abre, su contenido sale disparado empapándolo sin remisión. ¿A qué obedece ese fenómeno? A que la agitación produce una turbulencia tan intensa en el líquido que algunas bolsas pequeñas de gas se escapan de la solución en las zonas donde la turbulencia reduce la presión local.

Al abrir una botella de cerveza que no se haya agitado, el anhídrido carbónico sólo dispone de un sitio importante por el que salir de la solución: la superficie libre de la cerveza. Pero las cosas ocurren de otro modo en el caso de un envase de cerveza agitado: se generan cientos, y hasta millares, de diminutas burbujas, cada una de las cuales ofrece una superficie por donde el anhídrido carbónico sobrante puede huir de la solución. Entonces, al destapar el envase, el paso del anhídrido carbónico de disuelto a no disuelto es tan veloz que se produce la expulsión de líquido.

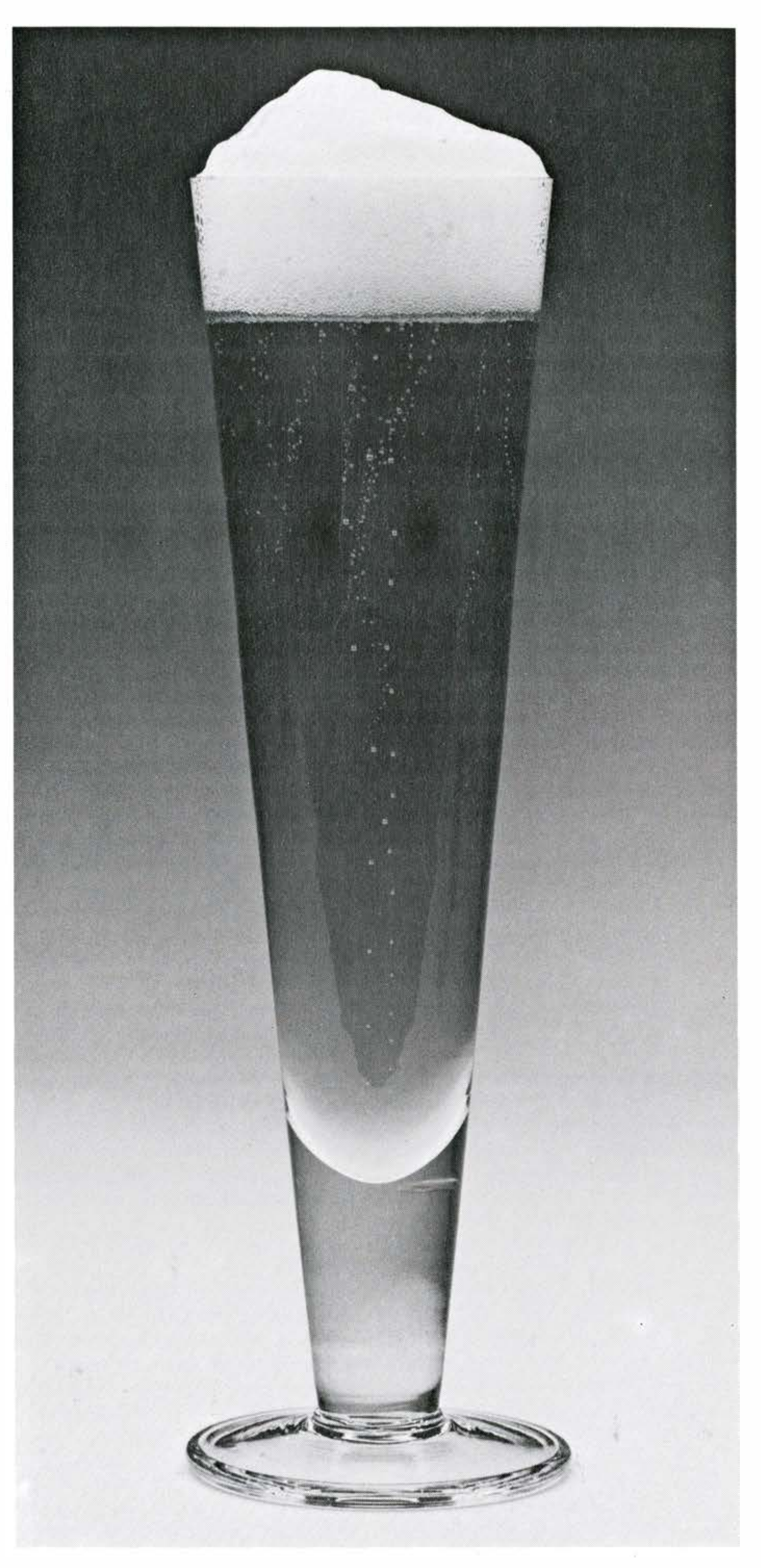
En un vaso de cerveza sin espuma, el anhídrido carbónico sale más pausado de la solución. Aquí las burbujas se forman, se desprenden de la pared del vaso y suben a la superficie, donde acaban reventando. Las burbujas se generan todas en la pared o en el fondo, si bien lo hacen exclusivamente en las irregularidades de la superficie del vidrio o donde se haya alojado una mota de polvo. Ello se debe a que, en ausencia de cavitación, las burbujas necesitan núcleos de condensación. Pero no son los núcleos de condensación en sí los que engendran las burbujas, sino el aire y el anhídrido carbónico que se absorbe en la superficie de los mismos.

Imaginemos que hubiera una pequeña grieta en el interior del vaso. En ella podría alojarse aire cuando lo llenáramos de cerveza. Dicha bolsa de aire ofrecería una superficie por donde escapar de la solución el anhídrido carbónico. Empezaría entonces a desarrollarse una burbuja de anhídrido carbónico, que acabará haciéndose suficientemente grande para desprenderse y salir a la superficie. Por otro lado, es probable que en la grieta quede aire y anhídrido carbónico suficientes para estimular la formación de otra burbuja.

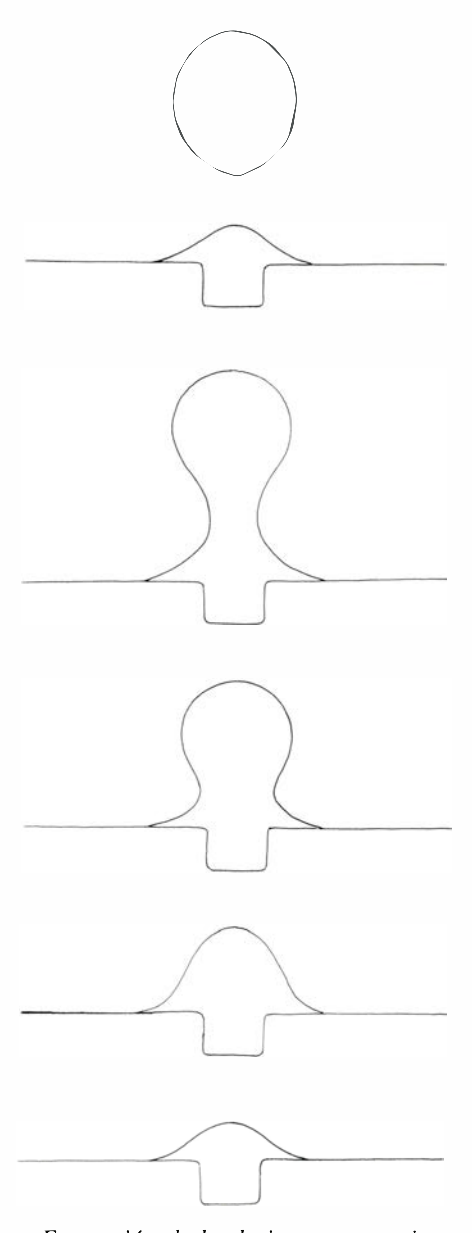
Una mota de polvo ejerce el mismo efecto si en su superficie hay aire absorbido. En una cerveza, como en cualquier otro refresco, puede fomentarse la formación de burbujas añadiendo una superficie sólida sobre la que pueda absorberse aire. Quizás el lector ya haya advertido que al introducir cubitos de hielo en un refresco carbónico aumenta el ritmo de producción de burbujas, con la efervescencia consiguiente. Lo que sucede en este caso es que en la superficie del hielo permanecen atrapadas minúsculas burbujas de aire y, cuando se funde dicha superficie, quedan al descubierto las bolsas de aire; éstas actúan entonces de núcleos, o franquicia, por donde escapará de la solución el anhídrido carbónico.

Toda burbuja formada a partir del aire contenido en una grieta de un vaso lleno de cerveza no es nunca esférica. A medida que en la burbuja penetra más anhídrido carbónico, crece la misma y se alarga hacia arriba. Además, cambia de forma. A medida que dilata su volumen, la superficie de arriba permanece más o menos semiesférica; la porción inferior, en cambio, se estrecha formando un cilindro. Se trata de un cilindro inestable: si la burbuja prosigue en su expansión y el cilindro continúa estrechándose, llegará un momento en que éste acabe por desinflarse, escapándose así la burbuja.

En la formación de una burbuja a partir del aire contenido en una grieta intervienen dos fuerzas. Por una parte, como el gas es menos denso que la cerveza, el empuje de Arquímedes lo impulsa hacia arriba, separándolo de la grieta. Por otra, la fuerza que contrarresta el empuje de Arquímedes es, al menos inicialmente, la tensión superficial existente en la superficie de separación entre la cerveza y el gas. Esta fuerza tiene su origen en la atracción mutua entre las moléculas de líquido que forman dicha superficie, y su efecto es la tensión que sufre la misma. En todos los puntos de la superficie curva de separación, la fuerza debida a la tensión



Nacimiento de burbujas en un vaso de cerveza alto



Formación de burbujas en una grieta

superficial es perpendicular a la superficie. Inmediatamente después de que la burbuja comienza a formarse, esa fuerza es un vector que apunta hacia el centro de la burbuja, o sea, hacia abajo.

La burbuja, al desarrollarse, hace que aumente el empuje de Arquímedes y éste acaba venciendo la tensión superficial. Ambas fuerzas opuestas no entablan un simple tira y afloja en sentidos contrarios. El empuje tira de la burbuja cada vez más, alargándola a partir de la grieta, y así la parte inferior comienza a adquirir la forma de un cilindro con la superficie lateral cóncava. Cuando se ha estirado la burbuja hacia arriba lo suficiente, la tensión superficial del cilindro ya no puede aguantar más la situación y el cilindro desaparece proyectando la parte superior de la burbuja. En ese momento, la tensión superficial confiere a la burbuja una forma algo más parecida a una esfera.

¿Por qué no se forman burbujas en los envases cerrados y en los abiertos dejan de formarse al cabo de una hora aproximadamente? En un envase cerrado, la presión relativamente elevada que reina en el líquido (debida a la que ejerce el gas encerrado en la parte superior) crea una presión elevada en el gas absorbido por los núcleos de formación; en éstos hay unas diminutas bolsas de gas donde entran y salen constantemente moléculas de anhídrido carbónico. Al estar la cerveza en equilibrio, el número de moléculas que entran por segundo es igual al de moléculas que salen, ya que la elevada presión existente en la bolsa de gas impide que en ella penetren moléculas adicionales que ensanchen la burbuja.

Pero una vez abierto el envase, disminuye la presión del gas, penetran más moléculas y la burbuja crece. Este proceso se interrumpe periódicamente cada vez que se desprende una burbuja; pese a ello, prosigue hasta que la reserva de moléculas de anhídrido carbónico se reduce sustancialmente. En este punto, el número de moléculas que penetra vuelve a igualarse al de moléculas que salen, y cesa el crecimiento de la burbuja.

El tamaño de las burbujas aumenta a medida que ascienden. La razón principal de ello estriba en que se desplazan hacia zonas de presión hidrostática cada vez menor, y pueden así dilatarse. Un buceador que haya pasado algún tiempo sumergido sufrirá las consecuencias de este fenómeno si emerge con demasiada rapidez. Lo que le ocurrirá es que, bajo presión, el nitrógeno contenido en el aire se le disolverá en la sangre; y, al remitir la presión, se le formarán burbujas en los fluidos orgánicos y padecerá los síntomas de la llamada enfermedad de los buzos. Con un ascenso más lento, el nitrógeno podrá ser reabsorbido por la sangre y eliminado en los pulmones.

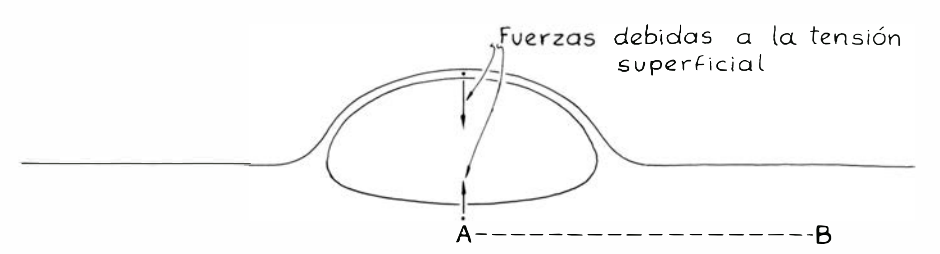
Cuando una burbuja llega a la superficie de la cerveza, perdura varios minutos antes de reventar. Esta burbuja no es esférica, ni mucho menos. La presión hidrostática subyacente, e inmediata, a la burbuja (sea éste el punto *A*) debe ser igual a la presión existente en todo punto situado a la misma altura y a gran distancia de la burbuja; de lo contrario, se organizaría una corriente de cerveza entre ambos puntos. En *B* la presión hidrostática resulta de la presión atmosférica y del peso de la cerveza que se encuentra sobre dicho punto. En *A* actúa la misma presión atmosférica, pero el peso del líquido existente encima es ínfimo porque allí sólo hay una delgada película de líquido. ¿Cómo pueden ser entonces iguales las presiones en *A* y *B*?

La presión adicional en *A* la proporciona la tensión superficial de la bóveda arqueada que forma la parte superior de la burbuja. La curvatura de esta bóveda contribuye con una fuerza dirigida hacia abajo. La tensión superficial del fondo cóncavo de la burbuja contribuye con una fuerza dirigida hacia arriba. El fondo no es, empero, tan curvo como la parte superior y la fuerza dirigida hacia arriba será, pues, menor. La forma de una burbuja flotante da origen a una fuerza resultante dirigida hacia abajo, lo que determina que la presión inmediatamente debajo de la burbuja aumente hasta el valor de equilibrio.

Si la burbuja llega a la superficie junto a la pared del recipiente, se avanza hacia ella y se le adhiere. Si hay otra burbuja cercana, se aproximan mutuamente. La atracción entre dos burbujas, o entre una burbuja y la pared, se debe a la desigualdad de las presiones existentes en las superficies que rodean a una burbuja. Entre dos burbujas cercanas la superficie de líquido es curva, no plana como en otras direcciones. Se trata de una curva cóncava, lo que significa que la fuerza debida a la tensión superficial empujará hacia arriba y, en consecuencia, disminuirá la presión en el seno de la cerveza situada cabe la superficie curva. En las demás zonas de las burbujas no remite la presión, porque la superficie de líquido es plana. En resumen, las burbujas salen al encuentro impulsadas por la desigualdad de las presiones.

Cuando una burbuja se encuentra próxima a la pared, la superficie de líquido comprendida entre la burbuja y la pared es más curva que la superficie del líquido situado en el lado opuesto de la burbuja. Vuelve a darse aquí que la curvatura genera una tensión superficial que reduce la presión en la zona superficial del líquido; la burbuja se ve impelida hacia la pared.

Si dos burbujas idénticas se tocan, su pared común, llamada lamela, es plana: a ambos lados de la misma las presiones de gas son iguales. Si las burbujas no son del mismo tamaño, la lamela se curva para que se originen presiones iguales en ambas caras. En este caso, las presiones son inicialmente distintas porque difieren las tensiones superficiales de las burbujas. La burbuja mayor, cuyo radio de curvatura es mayor, tiene una presión interna me-



Forma de una burbuja en la superficie de un vaso de cerveza

nor porque la tensión superficial de su pared es menor también. Así, cuando una burbuja grande entre en contacto con otra más pequeña, la pared común a ambas se adentrará en la burbuja mayor, de modo que la variación de tensión superficial debida a tal abombamiento nivele las presiones en ambas. (En una espuma de burbujas de múltiples tamaños diferentes no es posible una igualación absoluta de las presiones.)

En una espuma ideal cuyas burbujas tuvieran todas el mismo tamaño, las burbujas se organizarían en grupos de tres, en contacto unas con otras; y sus lamelas se intersecarían simétricamente en un punto llamado límite de Plateau (en honor de Joseph A. F. Plateau, quien estudió las pompas de jabón en el siglo XIX). Esta organización es la más estable frente a las sacudidas.

Aun cuando una espuma de cerveza estuviera formada por burbujas iguales y organizadas del modo descrito, se vendría abajo a los pocos minutos. Ni es ajena a ello la gravedad, que hace salir líquido de entre las lamelas, adelgazándolas hasta que su espesor queda sólo entre 20 y 200 nanómetros. Luego se estabilizan durante un rato, hasta que alguna sacudida fortuita procedente del entorno las mengua hasta unos cinco nanómetros, momento en que revientan.

La salida de líquido viene favorecida por una curiosa interacción de tensiones superficiales en el interior de las lamelas. Entre cada dos burbujas las superficies de la lamela son relativamente planas, de baja tensión superficial por tanto. En el lugar donde concurren tres lamelas para formar un límite de Plateau, empero, las superficies son curvas, las tensiones superficiales correspondientes son mayores y la presión en el seno de la zona de intersección es distinta de la que reina en la zona plana. Por ser convexas las superficies contiguas a la intersección, las fuerzas debidas a sus tensiones superficiales actúan alejándose de la intersección, reduciendo así la presión en el seno de la misma. La presión en el límite de Plateau resulta inferior a la existente en cualquier otro punto de las lamelas, con el efecto de que se aspira líquido desde las lamelas hacia el límite de Plateau y éstas pierden grosor.

La disminución paulatina del número de burbujas de una espuma resulta, además, de la difusión del gas que encierran a través de las delgadas paredes de las burbujas. Cuando una burbuja pequeña entra en contacto con otra mayor y no pueden igualar sus presiones en la pared común, la mayor presión reinante en la burbuja pequeña impulsa a parte del gas a disolverse en la pared. Este gas se difunde entonces y penetra en la burbuja mayor. Y es así como la burbuja grande se come a la chica.

En la resistencia de una espuma tiene mucho que ver la viscosidad superficial de las lamelas. No ha de confundirse la viscosidad superficial con la volúmica, que es la viscosidad del líquido comprendido entre ambas superficies. La viscosidad de las superficies es tan elevada que podrían adquirir una consistencia casi rígida, al menos hasta que las tensiones y esfuerzos cortantes que sufren alcancen determinado valor crítico, en cuyo instante cederán y la burbuja reventará.

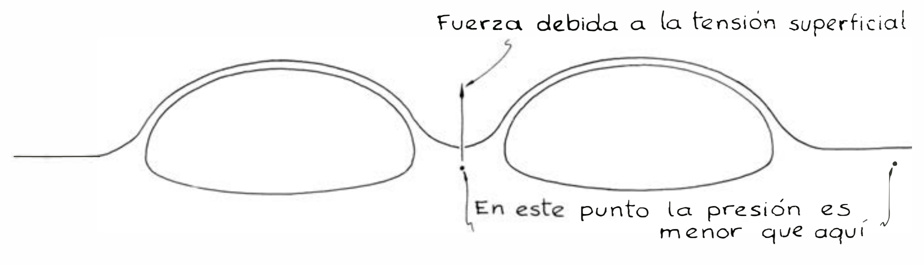
Hay espumas en las que fuerzas eléctricas contribuyen a consolidar las burbujas. En efecto, en cada una de las superficies de la lamela que separa dos burbujas contiguas puede encontrarse una capa de moléculas cargadas; entonces, cuando la lamela se adelgaza al perder líquido, las capas cargadas se aproximan entre sí y su repulsión mutua impide, o retrasa, la desaparición de la lamela.

Aunque tales capas dobles de cargas eléctricas sean importantes en ciertos tipos de espuma, ignoro si desempeñan un papel decisivo en la estabilización de la espuma de cerveza. En el caso de la cerveza no se conocen todavía bien los factores estabilizantes. Todo indica que las proteínas intervienen, así como los lúpulos curados. Otros ingredientes, verbigracia, sales de cobalto, hierro, níquel y cinc, se añaden para prolongar la vida de la masa espumosa o para reforzar la adherencia de las burbujas pequeñas a la zona de cristal que queda por encima de la espuma cuando se inclina el vaso.

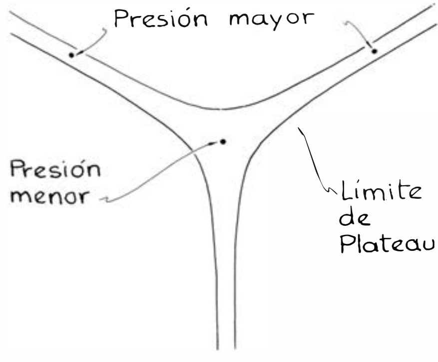
Aunque me viene de lejos la afición a las burbujas de cerveza (por motivos científicos, desde luego), no había caído en la cuenta de cierto fenómeno hasta que Craig Cook, alumno de la Universidad estatal de Cleveland, me lo mostró. Puede apreciarse en los vasos para cerveza Pilsner, tan estimados por los bebedores entendidos. Los vasos Pilsner son bastante anchos por la boca y se van estrechando hasta formar apenas una pequeña concavidad en el fondo. Para facilitar su manejo, el extremo cónico está inmerso en un grueso cilindro de vidrio. La singularidad de los vasos Pilsner reside en que la principal producción de burbujas se da en el fondo estrecho y cónico, donde puede verse el nacimiento de una corriente uniforme de burbujas, ninguna de las cuales tiene inicialmente un diámetro superior a unos 0,5 milímetros. En las paredes se forman burbujas en número mucho menor.

Las paredes se presentan perfectamente pulidas, mientras que el fondo parece áspero y lleno de diminutas crestas y grietas. Si uno de estos vasos está limpio y seco antes de recibir la cerveza, en las grietas quedarán atrapadas bolsas de aire, donde germinarán burbujas de anhídrido carbónico. Las burbujas se desprenderán en seguida de las grietas, y comenzarán su ascenso cuando todavía son pequeñas. Tan rápida producción de diminutas burbujas originará una corriente que se elevará hasta el centro de la superficie de la cerveza.

Es característico de las burbujas que nacen en las paredes del vaso que sean



Distribución de presiones en torno a dos burbujas



Lamelas que se forman en la reunión de tres burbujas

mucho mayores que las nacidas en el fondo. ¿Por qué? Entre otras razones, porque las burbujas del fondo sufren una presión hidrostática mayor y, en consecuencia, han de ser más pequeñas. Yo veo otro motivo en el grado de rugosidad. Hay en el fondo numerosos lugares contiguos donde pueden nacer burbujas, aunque ninguno de ellos sea capaz de crear una base suficientemente ancha para generar una burbuja grande. Además, la turbulencia que generan las burbujas cercanas al desprenderse de los puntos de nucleación limita asimismo el tamaño de las burbujas.

No me fue posible encontrar burbujas que se originaran en las paredes; cabe pensar que a causa de falta de bolsas de aire. Cuando sumergí en un vaso de cerveza un cuchillo, comenzaron a formarse burbujas sobre la hoja inmediatamente; pasado un rato, disminuyó la actividad. Para comprobar si aquella disminución se debía a la pérdida de anhídrido carbónico disuelto en la cerveza o a la pérdida de bolsas de aire en el cuchillo, introduje éste aún más en el líquido; entonces, aparecieron burbujas nuevas en gran número sobre la porción de hoja recién sumergida. Parece ser que el número de bolsas de aire nucleadoras decrece tras algunos minutos de formación de burbujas.

Pude observar que ocurría lo mismo en el vidrio. En un vaso limpio y seco las burbujas se generaban a ritmo vertiginoso. Pero el ritmo disminuía a cada nuevo recipiente de cerveza que vertía en el vaso. La acción de verter más cerveza generaba espuma y gran cantidad de burbujas, pero transcurrido un tiempo después del vertido la producción de burbujas caía. Evidentemente, el uso continuado del vaso reducía el número de bolsas de aire nucleadoras en sus paredes.

Con un cuchillo sumergido en cerveza acometí algunos ensayos más. Si inclinaba el cuchillo, se formaban capas de burbujas en la cara inferior de la hoja, al no hallar vía libre para escapar. Si le daba la vuelta, comenzaban a desprenderse de la hoja las capas de burbujas, evidenciando que la salida de una burbuja podría forzar la de otras.

La temperatura del cuchillo no parecía influir, pero sí su estado de limpieza. Cubrí el extremo inferior de un cuchillo con aceite de maíz y lo introduje en cerveza hasta más allá de la zona untada de aceite, con el resultado de que en ésta las burbujas se reunían en menor número que en la zona limpia, dado que el aceite tapaba las bolsas de aire absorbidas en el metal.

Donal Deneck, de Nueva York, me mostró en cierta ocasión cómo beber de una "yarda de cerveza", recipiente que mide entre dos y tres pies de altura, o sea, entre unos sesenta y noventa centímetros (tres pies hacen una yarda), y dotado de un fondo semiesférico y un cuello largo y de forma cónica cuya parte más estrecha está en el centro. Consumida la cerveza del cuello, cuesta bastante seguir bebiendo de ese recipiente. Para conseguir que la cerveza salga debe inclinarse el recipiente hasta que el extremo semiesférico esté más alto que la boca propia; pero si la inclinación es excesiva, en el cuello penetra una burbuja de aire de gran tamaño, deslizándose hacia el fondo, entonces elevado, y desplazando cerveza suficiente para procurar un pequeño remojón al bebedor.

¿Cómo se las ingenia Deneck? Inclina primero el recipiente hasta tener el fondo sólo un poco más alto que la boca. Luego golpea suavemente el costado del cuello para enviar, cuello arriba y hacia el fondo semiesférico, unas pocas burbujas de aire pequeñas, o parte de la espuma. Esas burbujas desplazan una cantidad razonable de cerveza. Así, corrigiendo la inclinación del recipiente y con golpecitos en el cuello del mismo, se las arregla Deneck para conseguir su ración de cerveza sin echársela por encima.

He recibido numerosas cartas relativas a la exposición que les hice en julio en torno a los fosfenos, esas figuras luminosas que podemos contemplar sin ayuda de la luz. Varios lectores me han descrito las imágenes de fosfenos que avisan de la aparición de una migraña y que la acompañan. Carol McAlpine, de Stroudsburg, me dice que, cuando le comienza una migraña, ve la red de vasos sanguíneos de la retina durante unos instantes, tras cerrar los ojos. A veces, la red vuelve a aparecérsele acompasada con el ritmo con su pulso. En plena crisis de migraña, ve ya "explosiones de oleadas nubosas en violeta, azul y gris, que flotan en todas direcciones y que cambian de forma continuamente". A medida que el dolor de cabeza se calma, las imágenes se convierten en formaciones de puntos del tipo que les describí con relación a la visión monocular de fosfenos.

Paul Tobias, de Los Angeles, me ha remitido un informe confeccionado por él y J. P. Meehan acerca de las imágenes de fosfenos provocadas por aceleraciones rápidas. Utilizaron voluntarios con los ojos vendados y asegurados con correas en una centrífuga, que se hizo girar en círculo a velocidad creciente. Cuando la aceleración efectiva alcanzaba un determinado valor umbral, los sujetos veían una distribución de puntos y estrellas azules. Para una aceleración efectiva de 3,6 g aproximadamente (unas 3,6 veces la aceleración de la gravedad), la distribución originaba una suerte de "gusanos dorados". Si el valor de la aceleración alcanzaba 4,5 g, los gusanos formaban un dibujo geométrico de color amarillo-naranja brillante. Sobre este dibujo había quien observaba una pulsación.

Se paraba la centrífuga. Pero no se les quitaba inmediatamente el vendaje: había que estudiar las imágenes secundarias de los fosfenos. La imagen secundaria, que perduraba hasta 90 segundos, consistía en una rosquilla o algo parecido a un eclipse total de sol sobre un fondo oscuro.

Esas imágenes aparecían sólo cuando la fuerza efectiva actuaba a lo largo del eje longitudinal del cuerpo; si actuaba de pecho a espalda, no se presentaban. Tobias y Meehan llegaron a la conclusión de que los fosfenos dependen de la rapidez de variación de la aceleración. Cuando ésta mantiene la dirección del eje longitudinal del cuerpo y varía con rapidez suficiente, lo que probablemente ocurra es que se reduzca la presión sanguínea en la retina y el ojo se relaje. Ambos efectos contribuyen a crear fosfenos.

Astronautas y pilotos podrían ver este tipo de fosfenos durante los momentos de aceleración rápida, si la fuerza efectiva actuase a lo largo del eje longitudinal de sus cuerpos. En realidad podría contemplarlos cualquier individuo sobre una instalación de rotación rápida de un parque de atracciones. En ciertos viajes de esa clase se desarrollan aceleraciones superiores a 3g y cualquier viajero, de suficiente presencia de ánimo, podría observar los fosfenos a esa velocidad. Por mi parte, he de confesar que, en esos momentos, me preocupa más mi integridad física.

Libros

Distintas maneras de narrar la historia de la ciencia y de interpretar esa labor

Ramón Margalef, Luis Alonso, Agustin Albarracín y F. G. Valdecasas

THE EVOLUTIONARY SYNTHESIS. PERSPECTIVES ON THE UNIFICATION OF BIOLOGY, publicado por Ernst Mayr y William B. Provine. Harvard University Press, Cambridge, Massachusetts, y Londres, Inglaterra, 1980; 487 páginas. Este volumen es resultado de una conferencia, celebrada en dos sesiones, en mayo y octubre de 1974, y organizada por el Comité de Historia Reciente de la Ciencia y la Tecnología, de la Academia Americana de Ciencias y Artes. Dada la índole del tema, el retraso en la publicación no es grave.

El objetivo central, al parecer, era examinar la convergencia de distintas ciencias biológicas en torno a lo que suele denominarse teoría sintética de la evolución, o neo-darwinismo. La conferencia debió ser única, por representar la última oportunidad de reunir, personalmente, o por sus contribuciones escritas, a varios nombres venerables. Algunos de ellos, como T. Dobzhansky, I. M. Lerner, E. Boesiger, B. L. Astaurov, fallecieron antes de publicarse este libro; otros no habían podido asistir personalmente a la reunión (B. Rensch y G. G. Simpson). La introducción ha sido escrita por Mayr que, como es de suponer, echa un poco de agua hacia el molino de la biogeografía y de la especiación alopátrica. El texto está distribuido en tres secciones: una por temas (genética, citología, embriología, sistemática, botánica, paleontología y morfología); otra por países (Unión Soviética, Alemania, Francia –"France is a kind of living fossil in the rejection of modern evolutionary theory" (sic)–, Inglaterra y Estados Unidos); la tercera parte incluye algunas consideraciones finales y notas biográficas.

La mayor parte del texto, en lo que se refiere a historia de la ciencia y a la presentación de datos, reúne materiales bien conocidos, pero se lee con gusto, principalmente por las notas biográficas o autobiográficas intercaladas que dan considerable vivacidad y ayudan a entender ciertos hechos. Probablemente una interpretación más seria requería el punto de vista del historiador y la colocación dentro de un contexto cultura más amplio. Sin esto, muchas páginas quedan en la anécdota y el chismorreo entre viejas glorias. Se destaca con particular énfasis la contribución del trío Fisher, Haldane y Wright, al que se añade –y con razón– Chetverikov. Precisamente, los relatos sobre la biología en Rusia, antes y después de la Revolución de 1917, de Adams y Dobzhansky, ilustran episodios poco divulgados en las presentaciones de historia de la biología usuales en el Oeste. Los participantes más jóvenes, como S. J. Gould y R. C. Lewontin, representan el contrapunto al desfile de los clásicos. Ambos tienen cosas interesantes que decir.

La mayor parte del texto muestra un grado de complacencia que no tiene razones muy profundas. La posible tautología de las ideas de fitness y selección apenas se menciona más que como una crítica (por supuesto injusta) a Morgan, por habérsele ocurrido en 1914. No se da una razón satisfactoria ni una simple mención de la existencia de otros problemas, como es la historia de las estirpes y su descripción. Quizá por esto no fue posible reclutar un consenso más amplio de los paleontólogos, y Simpson, a pesar de su docilidad dentro del neo-darwinismo, probablemente no quedó contento con lo que refleja este volumen.

El libro tiene, en conjunto, un carácter anticlimácico. Muchos lectores se preguntarán, como yo me pregunto, en qué consiste esta síntesis para que pueda ser considerada como un acontecimiento de suma distinción, que merezca una atención especial, y no como algo corriente en el quehacer científico. Porque toda ciencia, y especialmente una ciencia tan amplia como la biología, nos aparece como un haz de direcciones de trabajo, más o menos independientes unas de otras, por lo menos en su origen. En su avance se ejercen un saludable efecto recíproco de poda y depuración, al abandonarse las hipótesis y los quehaceres no respaldados por las otras disciplinas paralelas, y, al final, se manifiesta una aproximación obvia y casi total entre las direcciones compatibles que forman el nuevo eje coherente del avance. La síntesis evolucionaria no es más –ni menos y ello es importante– que esto, algo muy corriente en el progreso de cualquier rama de la ciencia. (R. M.).

HISTORIA CRÍTICA DEL PENSAMIENTO ESPAÑOL. Tomo III. Del Barroco a la Ilustración (siglos XVII y XVIII), por Jose Luis Abellán. Espasa-Calpe, S.A.; Madrid, 1981. Están brotando como hongos los aficionados a la historia de la ciencia en España. No tendríamos más que plácemes si los interesados supieran frenar a tiempo sus impulsos de llevar a la imprenta opiniones no contrastadas, refritos indigestos y tópicas descripciones sombrías sobre nuestra paramera científica desde que el movimiento tectónico nos ha individualizado geofísicamente. Al no existir institucionalizados los estudios de historia de la ciencia, si exceptuamos el caso de medicina, farmacia y el recentísimo de historia de la ciencia y la tecnología en la Universidad de Barcelona, algunas facultades de filosofía han optado por "completar" el título de sus enseñanzas de la siguiente guisa: "departamento de historia de la filosofía y la ciencia", "historia de la filosofía y la ciencia antiguas", id de las medievales, modernas y contemporáneas. Cambio o añadidura que no ha comportado ninguna incorporación nueva de personal preparado. Y, así, el docente obligado a explicar Aristóteles habrá de enseñar, "además", la embriología, la sistemática y la fisiología animal de los *Parva naturalia*; quien comente las críticas kantianas tendrá que explayarse, "además", en la astronomía de Laplace y del solitario de Konisberg; terminado el programa sobre la filosofía de Popper y de Quine, el profesor se aprestará a desentrañar las geometrías no-euclidianas y la asimetría entre materia y antimateria.

Evidentemente, no es ese el camino metodológico idóneo. La vía opuesta –del paradigma científico al filosófico– sería más ajustada a la realidad de la historia y al propio proceso de un pensamiento riguroso. Pero subyace todavía en nuestro subconsciente nacional la creencia –el prejuicio– de que el intelectual, el pensador, es el que domina los saberes humanísticos: desde la interpretación psicoanalítica del Quijote hasta los recursos estéticos de *Cien años de soledad*. Y se precian más de salir en el suplemento literario del *Times* que de conocer los materiales de

que consta su máquina de escribir. Cuando se ocupan de la ciencia, en sus escarceos superficiales, no es más que para repetir los lugares comunes del atraso, la insolidaridad entre los grupos y la intransigencia de los poderes obscurantistas. La ciencia en España necesita redimirse de quienes se postulan sus defensores. Nada más injusto que la ignorancia culpable.

Por eso, ahora que comienza a expandirse la voluntad de recuperar nuestro patrimonio científico, convendría revestirnos de una buena dosis de humildad y aprender de quienes van por delante en el camino a seguir. Por citar dos veneros de extraordinaria fecundidad: J. M. López Piñero y L. García Ballester en historia de la medicina y Juan Vernet en historia de la ciencia árabe. Hay bastantes más, por fortuna. ¿Por qué cito, pues, a esos? Porque personifican la formación necesaria para convertirse en buen investigador del pasado de una disciplina: conocimiento de la materia científica, dominio del idioma o de los idiomas en que están escritas las obras (clásicos y modernos) y una sólida cultura historiográfica que les permite encuadrar el nivel de una ciencia en el contexto académico, social y económico.

Quien no se ajuste a esos parámetros cometerá los desaguisados que encontramos, por ejemplo, en la obra que da pie a esta reseña. El tomo en cuestión cubre el período de gestación de la ciencia moderna. Una buena oportunidad para comparar el momento científico español –peninsular y ultramarino– con el del resto de Europa. Una oportunidad perdida.

Así, a propósito de la ciencia del Barroco. En la página 365 escribe sobre Martín Martínez: "En algún modo vinculado con el foco de Sevilla está el llamado DOCTOR MARTIN MARTINEZ (1684-1734), por la dedicatoria de su *Philosophia scéptica* a la 'Academia Hispalense, que en el solo espacio de seis lustros ha ilustrado más la Physica y Ciencias Naturales, que todas las demás Escuelas de España en algunos siglos' (*Prólogo*). Escribió varias obras de medicina, como *Noches anatómicas* (Madrid, 1716, 1750), *Anatomía completa del hombre* (Madrid, 1728), *Juicio final de la Astrología* (Madrid, s.f.). Su libro más famoso es el citado en primer lugar, y cuyo título completo es el siguiente: *Philosophia scéptica, extracto de la physica antigua, y moderna, recopilada en diálogos, entre un Aristotélico, Cartesiano, Gassendista, y Scéptico para instrucción de la curiosidad Española* (Madrid, 1720-1750). Su escepticismo le inclina más a Bacon que a Descartes..." Y como si no lo hubiera mencionado 43 páginas antes, vuelve a ocuparse del famoso médico en la 408: "Este es el caso del médico madrileño MARTIN MARTINEZ (murió en 1734), autor de una *Philosophia scéptica* (1730) y de una *Medicina scéptica* (1735). Aunque mantiene el ámbito de la fe marginado de sus especulaciones médicas y filosóficas, en éstas su escepticismo es bastante radical, si bien dentro de una fuerte base empírica. El libro donde desarrolla sus ideas al respecto es el mencionado anteriormente, en primer lugar, y cuyo título completo dice así: *Philosophia sceptica, extracto de la physica antigua y moderna, recopilada en diálogos entre un Aristotélico, Cartesiano, Gassendístico y Sceptico, para introducción de la curiosidad española* (Madrid, 1930." ¿Qué comentario de texto sacaría un bachiller? Diría, por lo menos, que el autor no ha manejado los libros originales que cita, ni se ha preocupado de compulsar qué fuente de inspiración es la acertada. Que ni siquiera ha repasado el texto transcrito (las fechas cambian de un párrafo a otro e incluso en el mismo párrafo, aunque esto último obedece a un mero error tipográfico). Hilando más fino, el alumno pondría en duda la fecha asignada a la *Medicina sceptica*, 1735, si Feijóo escribió un libro que llevaba por título *Aprobación apologética del scepticismo médico del doctor don Martín Martínez* (Oviedo, 1725), como recoge Abellán en la página 366. La verdad histórica parece inclinarse por datar ese libro entre 1722 y 1725. Y si al bachiller le cedemos un ejemplar de la *Philosophia Sceptica* hallará nuevas pruebas de cómo los fragmentos citados en la obra de Abellán han sufrido la erosión y transformación que caracteriza el recurso a los intermediarios. Y como parece fiarse en trabajo ajeno más que en la propia incursión, pasa por alto la física (mecánica, óptica y acústica) de Martínez y la influencia que en él tuvo, entre otros, Robert Boyle.

Si se me permitiera una frase facilona diría que Abellán se pierde en el barroquismo de su ilustración, de un ingente fichero de nombres, que le importan más que la doctrina o la aportación a la geología, a la geografía, a la física, a los procesos químicos. Del paso de Proust por España, ya en la parte dedicada a la Ilustración, dice que "realizó importantes experimentos para encontrar una pólvora de mayor calidad y perfeccionó las técnicas de fabricación de porcelanas y tintes". Nada de su formulación de la ley de las proporciones definidas. Y siguiendo con la química, la oposición de Aréjula a Lavoisier, con el tras-

fondo de la renovación de la nomenclatura no merece su atención. No busque el lector luz sobre la introducción del sistema linneano en España, ni la oposición al mismo por parte de José Quer (de quien por cierto ni siquiera cita que constituye el primer apologista de la ciencia española, tema redundante en este volumen de Abellán). Curiosa es la interpretación que da de Félix de Azara, un naturalista excepcional necesitado de una buena biografía (que ponga al día la de Alvarez López y ahonde la de Beddall). Tras insistir en su visión anticipada de la evolución de las especies, dice que "sus puntos de vista fueron aceptados por Cuvier y Darwin". George Cuvier era fijista y Darwin evolucionista. ¿Cómo pudo influir en ambos? Habría que haber expuesto que la influencia se ejerció, obviamente, en otros campos de la biología; en el de la anatomía comparada en el primero y en el gradualismo evolutivo en el segundo. Si alguien quiere tener una idea más exacta de qué nivel alcanzó la técnica en la ilustración acuda, por ejemplo, a *Ciencia y tecnología en la España ilustrada*, de Antonio Rumeu de Armas (Madrid, 1980), libro que aquí no se cita.

Es José Luis Abellán un destacado especialista en historia de la filosofía española, con obras sobre Unamuno, Ortega, el erasmismo, etcétera. Bienvenido sea su afán por salvar ese vado que separa la ciencia y la filosofía en España. Pero la historia de la ciencia no puede vivir de buena voluntad, a modo de apéndice residual de la expresión filosófica y literaria. Abandonemos de una vez por todas la retórica –sea diatriba sea panegírico– y comencemos a usar regla y compás, a computar en números romanos, a reconstruir los laboratorios alquimistas, a ponernos en la piel del medieval o del moderno, para entender el mundo tal como ellos lo entendieron. Aquí, menos que en ninguna otra profesión, el hábito no hace al monje: no por llamarnos profesores de historia de la filosofía y la ciencia vamos a ser maestros en lo segundo. (L. A.)

RAMÓN Y CAJAL, por Dorothy F. Cannon. 4.ª edición. Ediciones Grijalbo, S. A.; Barcelona-México, 1981; 254 páginas. Dorothy F. Cannon, especialista en lenguas romances formada en el Conneticut College de New London, directora un día de *El Farmacéutico*, revista médica y farmacéutica editada en español, publicó en 1949 un libro, *Explorer of the human brain. The life of Santiago Ramón y Cajal (1852-1934)*, editado por Henry Schuman Inc. de New York. Mostraba en él la autora su admiración por la vida y la figura del sabio español, vida y figura que ofrecía al público de lengua inglesa a través de una versión propia, en la que se entremezclaban, bajo el cañamazo de la utilización de los *Recuerdos de mi vida* de Ramón y Cajal, una serie de testimonios de autores que conocieron su obra e incluso trataron a su persona, y unos breves capítulos históricos en torno a los principales eventos de la vida española coetánea de Cajal. De otra parte, le había sido fácil acceder a la autobiografía cajaliana a través de la traducción al inglés que de ella había llevado a cabo en 1937 el profesor E. H. Horne, de la Universidad canadiense de Toronto. Prestaba especial interés al librito un sabroso prólogo del fisiólogo inglés Charles Scott Sherrington, quien en breves páginas subrayaba rasgos y anécdotas inéditas de nuestro premio Nobel, directamente aprehendidas durante la permanencia de Cajal en Londres, huésped personal del profesor británico, anécdotas acompañadas de unas reflexiones sobre el significado de los descubrimientos de aquél, tanto en su aspecto morfológico como fisiológico, este último, es bien sabido, de tanta trascendencia para la obra ulterior de Sherrington.

Aparecía el libro de Dorothy F. Cannon en 1949. Habían transcurrido quince años desde la muerte de Ramón y Cajal, y sólo una gavilla de artículos necrológicos, por lo general fechados en torno a 1934, año de su óbito, permitían conocer al gran público, casi siempre muy parcialmente, su obra y el significado de la misma. Hágase excepción, en tal sentido, del trabajo de Tello titulado *Cajal y su labor histológica*.

Era indudable, por tanto, la oportunidad del libro, que dos años después fue traducido al castellano en México, con una introducción de J. F. Tello. ¿Acertó también la Sra. Cannon en el contenido de su aportación a la bibliografía cajaliana? Muy pocos meses antes de que las prensas alumbraran este libro había aparecido en España el primero de una serie de trabajos y estudios de Pedro Laín Entralgo consagrados a la exposición del significado de la obra de Cajal, dentro de un libro cuyo título general rezaba *Vestigios. Ensayos de crítica y amistad*. Mostraba Laín en él cómo la simple historia externa no bastaba para conocer realmente el sentido y las consecuencias de la hazaña cajaliana, precisa de una interpretación y comprensión profundas, capaces de desvelar el significado total de lo que hizo y del modo como llegó a hacerlo la persona de Cajal.

Pero el libro de Dorothy F. Cannon se limitaba, repito, a utilizar el propio texto de la *Autobiografía* del aragonés, a resumirlo y a adornar tal resumen con la reproducción de párrafos significativos del propio Cajal; todo lo cual venía a completarse intercalando en el contexto breves capítulos históricos –El reinado de Isabel II, Consecuencias de la revolución, Era de paz, La regencia, El krausismo y la Institución Libre de Enseñanza, La pérdida de las últimas colonias, Los primeros años de Alfonso XIII, Ultimos años del reinado de Alfonso, Dictadura en España, La Segunda República española–, otros histológicos –Un vistazo al sistema nervioso y El "tercer elemento"– así como el testimonio de algunos discípulos y científicos como Fernando de Castro, Wilder Penfield y Moisés Polak. Varias ilustraciones –un mapa del nordeste de España, esquemas histológicos de E. Home Craigie y J. F. Tello– y sucintas notas biográficas, a pie de página, de las personalidades citadas en el texto, daban remate a la obra, en sus dos versiones, inglesa y castellana.

Tras la lectura de este *Ramón y Cajal* acuden a la mente del lector insatisfecho una importante serie de cuestiones. ¿Dónde encontrar en él los orígenes de la vocación neurohistológica de Cajal? ¿Dónde la raíz y expresión de su patriotismo? ¿Dónde el significado histórico de sus descubrimientos? ¿Dónde, en fin, toda la problemática sociológica, cultural, económica, de la ciencia española del siglo XIX, así como los antecedentes de la escuela histológica española y los avatares de la misma, tras la muerte de Cajal? Al limitarse la autora al testimonio del propio don Santiago, es evidente que falta en su libro esa perspectiva que permite esclarecer las respuestas a las cuestiones aludidas. D. F. Cannon nos ofrece una historia externa y parcial de la biografía de Cajal –en el sentido de mera crónica o narración de su vida, apoyada en el pensamiento y en el sentimiento del autor de esa vida– y no una reactualización del pasado como nos pedían Collingwood y Ortega. No voy a entrar, por supuesto, en la exposición de la siempre actual y polémica dicotomía entre la "historia interna" y la "historia externa", siquiera confiese mi inclinación por la solución de compromiso ofrecida por Canguilhem, en 1975, al distinguir entre el objeto de la historia de las ciencias y el objeto de las ciencias: creo, por tanto, en la necesidad de asumir el externalismo y el internalismo en una visión total de la historia de un problema científico o de una vida científica.

En 1949 era disculpable, desde los Estados Unidos, esta visión parcial de Dorothy F. Cannon, y su libro vino a cumplir, repito, un hito importante en la bibliografía cajaliana. Pero desde 1949 a 1981, fecha de la edición castellana que comento, ha sido muy amplia la tarea llevada a cabo por P. Laín Entralgo, Luis S. Granjel, J. M.ª López Piñero, García Durán y F. Alonso Burón, D. Ferrer, S. Loren, H. Williams y el autor de esta recensión, tendente al logro de esa visión asuntiva que antes proclamaba. Por supuesto que D. F. Cannon no la conoció ni estaba obligada, es evidente, a tenerla en cuenta. Pero pienso que, al menos, en la bibliografía general e histológica que al final del libro aparece, debían los editores haberla citado como apéndice, convirtiendo así este *Ramón y Cajal* que se queda en mera reproducción de un libro ya histórico, en el testimonio vivo y actual que nuestro sabio merecía y que la cuidadosa edición de Grijalbo exigía. No ha sido así: por tanto, cuanto aquí se apunta corresponde, no lo olvide el lector, a un libro aparecido hace treinta y dos años y cuyas referencias bibliográficas no traspasan la frontera del año 1945. (A. A.)

HUMBOLDT Y EL COSMOS: VIDA, OBRA Y VIAJES DE UN HOMBRE UNIVERSAL, por Douglas Botting. Ediciones del Serbal, S.A.; Barcelona, 1981. La vida del barón Friedrich Wilhelm Karl Heinrich Alexander von Humboldt se inscribe en uno de los períodos históricos más importantes de la cultura occidental; aquel que se extiende a caballo entre los siglos XVIII y XIX. Trascendentales sucesos tanto en el orden político como científico y cultural ocurren en el espacio relativamente breve para la historia de la humanidad como son los poco más de cien años que transcurren desde el comienzo de la publicación de la "Enciclopedia" hasta la capitulación de Sedan y la terrible experiencia de la Commune. Domina inicialmente el racionalismo en el sentido expuesto por Ferrater Mora, "la razón instrumento mediante el cual el hombre podrá disolver la obscuridad que le rodea", haciendo una deificación de la cultura y estimulando el conocimiento de la naturaleza, lo que despierta el entusiasmo por la ciencia: los nombres de los grandes naturalistas son populares en los salones de la aristocracia y clases acomodadas y los nuevos hechos físicos y químicos, motivo de regocijadas experiencias en ambientes sociales. Reyes y príncipes llaman a su lado y distinguen con su amistad a los grandes sabios, se discuten apasionadamente teorías y descubrimientos y, en fin, la vida intelectual y científica se infiltra en los antes frívolos círculos sociales. Esta época no ha sido bien definida, quizá por atenerse demasiado fielmente al concepto cronológico de los "siglos". El movimiento científico-cultural racionalista que se encuadra bajo el nombre de "la Ilustración" (Aufklärung) tiene como hecho más notable el de la "difusión" o dicho con las palabras de Julián Marías "por primera vez, escritores de carácter ideológico tienen un público muy amplio"; si se piensa que del *Teatro crítico universal* y de las *Cartas eruditas y curiosas*, del Padre Feijóo se vendieron 400.000 ejemplares en un país de apenas diez millones de habitantes, con enorme mayoría de campesinos iletrados, asombra la magnitud de la difusión. La característica científica fundamental de la Ilustración es la "interpretación visual del conocimiento" y sus armas van a ser la observación y la experimentación. No son del todo originales estos conceptos, pero sí es su aceptación universal y, sobre todo, como hemos dicho antes, su difusión popular. La consecuencia es el prestigio del "sabio" y su autoridad "una autoridad que puede compararse con la de la Iglesia" (palabras también de J. Marías). El conjunto es un movimiento intelectual jerarquizado. La naturaleza es la gran maestra universal y su dominio es la aspiración más deseable por el hombre: la ciencia tiene por misión el conocerla y el arte el imitarla. Por tanto, como la naturaleza se nos presenta como la más ordenada perfección, la actitud del hombre "racional" deberá ser adaptarse a las leyes naturales, al "dictado de la razón".

En contraste con este mundo de la razón, el movimiento cultural del Romanticismo parece representar todo lo contrario a la serenidad que ilumina el horizonte de la Ilustración. "Sensitivo, emocional, prefiriendo el color a la forma, lo exótico a lo familiar, ansioso de novedades, sediento de aventuras reales o imaginarias, gozándose en el desorden y en la incertidumbre, porfiando en la singularidad del individuo, hasta el punto de hacer de la excentricidad una virtud...." (The Encyclopedia of Philosophy, Collier Macmillan, New York-London, 1972). El típico romántico anhela la ruptura con las generaciones anteriores. El Romanticismo "era o creía ser en todas las cosas lo opuesto a la Ilustración" (Crane Brinton), pero con la perspectiva actual no podemos evitar el pensamiento de "que la rebelión romántica contra los padres era en sí mismo una prueba de la relación filial que les unía".

La estrecha relación entre Ilustración y Romanticismo, a pesar de su pretendida antinomia, lo vemos en las dificultades que se encuentran para poder clasificar las personalidades de la época, en especial las científicas, en uno de los dos movimientos culturales. Buffon, por la época y por su natural ser, es personaje indudable de la Ilustración. Pero, ¿y su seguidor Cuvier? ¿Pertenece a la Ilustración o al Romanticismo? ¿Puede negársele a Cuvier el culto a la Razón y la serenidad clásica en la sistematización? Y lo mismo que de Cuvier, podríamos decir de Lavoisier o de Laplace, de Cavendish o de Volta, de Priestley o de Watt y aún del propio Goethe. Entre todos, niguno quizá más representativo tanto de la Ilustración como del Romanticismo científico que Alexander von Humboldt. Caso excepcional de amor a la ciencia, metódico y sistemático en su trabajo –hacía mediciones hora a hora, día y noche, de la temperatura, la humedad, la presión barométrica, la declinación magnética; observaba, anotaba y experimentaba en los trópicos o en Siberia, en cualquier sitio en que estuviera– y al mismo tiempo sediento de aventuras, soñador, idealista...

La biografía de Alexander von Humboldt es apasionante. En sus tiempos el éxito de popularidad que obtuvo parece increíble, no tratándose de un actor de teatro, ni de un escritor ni de un político, sino de un científico exclusivo, cuyo amor e interés se cifraba sólo en la naturaleza y cuyo horizonte era sólo el conocimiento de todo lo que le rodeaba. Alexander era recibido tanto en París (que fue durante muchos años su residencia habitual) como en su patria o en los más alejados lugares en "olor de multitud", como si hubiera sido, valga la comparación, una estrella del cine o de la televisión. Una biografía de este ilustre hombre que dejó honda huella en la ciencia y en la aventura (su ascensión al Chimborazo tuvo caracteres de leyenda) ha de causar impacto en el lector. Podríamos decir que hacía falta esta biografía, que se echaba de menos. Porque habiendo tenido Humboldt tan enorme éxito de público en su tiempo y habiendo dejado tanta huella en la ciencia y la literatura, el gran público se ha desentendido de él. Que haya llegado a oídos del que esto escribe, en español (aparte de las citas habituales de las enciclopedias y libros de historia científica) sólo existe el breve capítulo biográfico correspondiente en la obra *Los descubridores célebres* de Leroi-Gourhan (Barcelona, Ed. G. Gili, 1964).

Douglas Botting escribe una biografía con estilo suelto y ágil (hay que alabar la excelente traducción), ateniéndose a la pauta cronológica, pero sin caer en la pesada relación de hechos. El análisis del ambiente que rodea al personaje desde su infancia en el seno de su hogar aristocrático y una clase social dominante, donde Alexander lo tenía todo resuelto para llegar a ocupar altos cargos bien remunerados, está perfectamente logrado, al igual que su carácter independiente e insatisfecho. Ama a la naturaleza y está decidido a sacrificarlo todo por ella. Parece increíble que este sentimiento lo tuviera desde la infancia y que a los veintiocho años rompiera con una situación estable y un puesto permanente en la administración, llevado por el sueño de exploraciones científicas en las lejanas tierras de las Antillas. En los momentos actuales no es posible darse cuenta de lo que esto significaba en aquella época. No se trataba tan sólo de los lentos viajes al paso de las caballerías o de las inciertas navegaciones a través del inmenso oceáno confiados en la fuerza del viento y también en que ésta no fuera demasiada. Se trataba, además, de las continuas guerras en una Europa agitada (no pudo llegar a Italia, no pudo ir a Túnez) y convulsa, dividida en multitud de países y estremecida por la mayor subversión ideológica de la historia. No le arredran los iniciales fracasos en su intento de estudiar los volcanes de Italia, ni la frustración de los primeros planes que se le ofrecen ilusos y románticos (expedición por el Nilo de un excéntrico mecenas inglés, vuelta al mundo concebida en la efervescente imaginación del París revolucionario). Botting nos conduce paso a paso por este deambular anhelante e ilusionado hasta alcanzar lo que sería la gran ocasión de su vida: su arribada a España. No encuentra aquí las dificultades que en otros países, recorre los campos con su sextante libremente y puede cartografiar el perfil de la península desde Valencia hasta La Coruña y desde Granada hasta el Pirineo. Consigue entrar en contacto con el Rey y despertar su entusiasmo. Obtiene de él fácilmente un pasaporte libre para viajar por toda la América española y garantía de que todas las autoridades y gobernadores le debían ayuda y protección. Emprende la gran epopeya que le permitiría saciar su ansia de conocimientos y de fama. Porque Humboldt, como muchos de los grandes investigadores, es vanidoso y teatral. Botting nos lo muestra celoso incluso de su gran amigo, y compañero permanente de expedición, Aimé Bonnpland. Y el lector, sobre todo el español, aprecia también esto en lo poco agradecido que se muestra no sólo por el permiso obtenido, sino porque en toda su expedición encuentra ayudas inestimables de españoles entusiastas y no sólo de la autoridad como el pasaporte real garantizaba. Así en la expedición a las fuentes del Orinoco y a comprobar la existencia del Casiquiare (que lo dudaran los sabios de París no quiere decir que los españoles no lo hubieran descubierto y recorrido) es acompañado fielmente por Nicolas de Soto y por las inhóspitas tierras va saltando de Misión en Misión, donde los sacerdotes, bien habituados a la tierra, le resuelven las muchas dificultades que podía encontrar. Lo mismo sucede en su expedición a los Andes, en la que el propio hijo del Gobernador, Carlos Montufar le acompañará durante todo el viaje y subirá con él al Chimborazo, aunque toda la gloria fuera para Humboldt.

La biografía está editada con gran cuidado. Profusamente ilustrada con grabados originales del propio Humboldt, ya de las obras publicadas, ya de manuscritos y cartas o del propio diario del biografiado, permiten al lector revivir las exploraciones, al mismo tiempo que sirven para comprobar la fidelidad con que se reproducía la imagen de las nuevas plantas y de las nuevas especies zoológicas descubiertas. Pero el libro contiene, además, otras numerosas ilustraciones de la época o fotografías actuales de los espacios recorridos y planos para facilitar la comprensión de los viajes y zonas exploradas.

En pocas palabras, este libro es la biografía apasionante de una personalidad científica, que durante los noventa años de su prolongada existencia cubrió casi todo un período de la historia, ocupando siempre un primer plano de actualidad, del que puede dar idea la enorme difusión de su última y más importante obra, "El Cosmos", cuyo primer tomo se publicó cuando ya tenía 76 años, el segundo cuando tenía setenta y ocho, el tercero cuando tenía ochenta y uno y el cuarto y último después de su muerte. El libro recoge una época que Humboldt vivió plenamente con sus numerosos amigos; científicos, humanistas, políticos y reyes, desde Goethe hasta Federico Guillermo IV de Prusia, con sus enormes progresos científicos, desde los grandes descubrimientos de la física con Volta, Galvani, y los grandes descubrimientos en la química (Lavoisier, Gay-Lussac), hasta coincidir su muerte con la publicación de una de las obras que más influjo han ejercido en la investigación de las ciencias de la naturaleza: *El origen de las Especies por selección natural*, de Darwin (F.G.V.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

ARMAS LASER

RESPONSE OF MATERIALS TO LASER RADIATION: A SHORT COURSE. J. T. Schriempf. Naval Research Laboratories, 1974.

HIGH ENERGY LASER WEAPONS: A TECHNICAL ASSESSMENT. M. Callaham y K. Tsipis. Program in Science and Technology for International Security Report No. 6.

NAVEGACION MAGNETICA EN LAS BACTERIAS

A NOVEL METHOD FOR THE ISOLATION AND STUDY OF A MAGNETOTACTIC BACTERIUM. T. T. Moench y W. A. Konetzka en *Archives of Microbiology*, vol. 119, págs. 203-212; 1978.

MAGNETITE IN FRESHWATER MAGNETOTACTIC BACTERIA. Richard B. Frankel, Richard P. Blakemore y Ralph S. Wolfe en *Science*, vol. 203, n.º 4387, págs. 1355-1356; 30 de noviembre de 1979.

ULTRASTRUCTURE OF A MAGNETOTACTIC SPIRILLUM. D. L. Balkwill, D. Maratea y R. P. Blakemore en *Journal of Bacteriology*, vol. 141, n.º 3, págs. 1399-1408; marzo, 1980.

SOUTH-SEEKING MAGNETOTACTIC BACTERIA IN THE SOUTHERN HEMISPHERE. R. P. Blakemore, R. B. Frankel y Ad. J. Kalmijn en *Nature*, vol. 286, n.º 5771, págs. 384-385; 24 de julio de 1980.

MAGNETOTACTIC BACTERIA AT THE GEOMAGNETIC EQUATOR. R. B. Frankel, R. P. Blakemore, F. F. Torres de Araujo, D. M. S. Esquivel y J. Danon en *Science*, vol. 212, n.º 4500, págs. 1269-1270; 12 de junio de 1981.

LIMITACIONES DE LA FIJACION BIOLOGICA DE NITROGENO

INCREASING CROP PRODUCTIVITY: THE PROBLEM, STRATEGIES, APPROACH AND SELECTED RATE-LIMITATIONS RELATED TO PHOTOSYNTHESIS. R. W. F. Hardy, U. D. Havelka y B. Quebedeaux en *Proceedings of the Fourth International Congress on Photosynthesis*, págs. 695-735. Editado por D. O. Hall, J. Coombs y T. W. Goodwin. The Biochemical Society, London, Colchester, 1977.

PHOTOSYNTHESIS AND PLANT PRODUCTIVITY. I. Zelitch en *Chemical and Engineering News*, n.º 57, págs. 28-48; 1979.

PHOTOSYNTEHSIS II. PHOTOSYNTHETIC CARBON METABOLISM AND RELATED PROCESSES en *Encyclopedia of Plant Physiology, New Series*, vol. 6. Editado por M. Gibbs y E. Latzko. Springer-Verlag, Berlín, Heidelberg, Nueva York; 1979.

LA FIJACIÓN BIOLÓGICA DEL NITRÓGENO. IMPORTANCIA Y PERSPECTIVAS. J. Olivares, E. J. Bedmar y J. Casadesús en *Anales de Edafología y Agrobiología*. En prensa.

JUPITER Y SATURNO

VOYAGER 1 ENCOUNTER WITH JUPITER. *Science*, vol. 204, n.º 4396; 1 de junio de 1979.

VOYAGER 2 ENCOUNTER WITH JUPITER. *Science*, vol. 206, n.º 4421; 23 de noviembre de 1979.

THE NEW SOLAR SYSTEM. Dirigido por J. Kelly Beatty, Brian O'Leary y Andrew Chaikin. Cambridge University Press y Sky Publishing Corporation, 1981.

VOYAGER 1 ENCOUNTER WITH SATURN. *Science*, vol. 212, n.º 4491; 10 de abril de 1981.

EL OIDO DE LA LECHUZA

ACOUSTIC LOCATION OF PREY BY BARN OWLS (*TYTO ALBA*). Roger S. Payne en *The Journal of Experimental Biology*, vol. 54, págs. 535-573; 1971.

HOW THE OWL TRACKS ITS PREY. Masakazu Konishi en *American Scientist*, vol. 61, n.º 4, págs. 414-424; julio-agosto de 1973.

SOUND LOCALIZATION IN BIRDS. E. I. Knudsen en *Comparative Studies of Hearing in Vertebrates*, dirigido por Arthur N. Popper y Richard R. Fay. Springer-Verlag, 1980.

FIBRINOGENO Y FIBRINA

THE AMINO ACID SEQUENCE OF THE α-CHAIN OF HUMAN FIBRINOGEN. R. F. Doolittle, K. W. K. Watt, B. A. Cottrell, D. D. Strong y M. Riley en *Nature*, vol. 280, n.º 5722, págs. 464-468; 9 de agosto de 1979.

TRINODULAR STRUCTURE OF FIBRINOGEN: CONFIRMATION BY BOTH SHADOWING AND NEGATIVE STAIN ELECTRON MICROSCOPY. Walter E. Fowler y Harold P. Erickson en *Journal of Molecular Biology*, vol. 134, n.º 2, págs. 241-249; 25 de octubre de 1979.

STUDIES ON SYNTHETIC PEPTIDES THAT BIND TO FIBRINOGEN AND PREVENT FIBRIN POLYMERIZATION: STRUCTURAL REQUIREMENTS, NUMBER OF BINDING SITES, AND SPECIES DIFFERENCES. Andrew P. Laudano y Russell F. Doolittle en *Biochemistry*, vol. 19, n.º 5, págs. 1013-1019; 4 de marzo de 1980.

ALGEBRA POR ORDENADOR

SYMBOLIC MATHEMATICAL COMPUTATION. David R. Stoutemyer y David Y. Y. Yun en *Encyclopedia of Computer Science and Technology. Vol. 15: Supplement*, dirigido por Jack Belzer, Albert G. Holzman y Allen Kent. Marcel Dekker, 1980.

PROCEEDINGS OF THE 1981 ACM SYMPOSIUM ON SYMBOLIC AND ALGEBRAIC COMPUTATION. Dirigido por Paul S. Wang. Association for Computer Machinery, 1981.

EL PRINCIPIO ANTROPICO

LARGE NUMBER COINCIDENCES AND THE ANTHROPIC PRINCIPLE IN COSMOLOGY. Brandon Carter en *Confrontation of Cosmological Theories with Observational Data: Proceedings of the Second Copernicus Symposium*, dirigido por M. A. Longair. D. Reidel Publishing Co., 1974.

ON WHAT GOD CHOSE: PERFECTION AND GOD'S FREEDOM. George Gale en *Studia Leibnitiana*, tomo 8, n.º 1, págs. 68-87, enero, 1976.

JUEGOS MATEMATICOS

WHAT WE KNOW AND DON'T KNOW ABOUT INFLATION. Robert M. Solow en *Technology Review*, vol. 81, n.º 3, págs. 30-46; diciembre/enero, 1979.

REAGANOMICS: SUPPLY-SIDE ECONOMICS IN ACTION. Bruce R. Bartlett. Arlington House Publishers, 1980.

TALLER Y LABORATORIO

SO WHAT'S A BUBBLE? Jeffrey C. May en *SciQuest*, vol. 52, n.º 8, págs. 16-20; octubre, 1979.

CLOUD PHYSICS IN A GLASS OF BEER. Craig F. Bohren y Gail M. Brown en *Weatherwise*, vol. 34, n.º 5, págs. 221-223; octubre, 1981.